OPEN ACCESS

University of BRISTOL

Tan, V., Biernacka, K., Dudding, T., Bonilla, C., Gilbert, B., Kaplan, R. C., Qi, Q., Teumer, A., Martin, R., Perks, C., the PRACTICAL Consortium, Timpson, N., & Holly, J. M. P. (2018). Reassessing the association between circulating Vitamin D and IGFBP-3: observational and Mendelian randomization estimates from independent sources. *Cancer Epidemiology, Biomarkers and Prevention*. https://doi.org/10.1158/1055-9965.EPI-18-0113

Peer reviewed version

Link to published version (if available):
10.1158/1055-9965.EPI-18-0113

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via AACR at https://doi.org/10.1161/CIRCULATIONAHA.117.033278 . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

Title: Reassessing the association between circulating Vitamin D and IGFBP-3: observational and Mendelian randomization estimates from independent sources.

Short title: The association between circulating Vitamin D and IGFBP-3

Authors:

Vanessa Y. Tan[1,3]*, Kalina M. Biernacka[2,6]*, Tom Dudding[1,3], Carolina Bonilla[1,3], Rebecca Gilbert[3], Robert C. Kaplan[4], Qi Qibin[4], Alexander Teumer[5], Richard M. Martin[1,3,6], Claire M. Perks[2], the PRACTICAL consortium[#], Nicholas J. Timpson[1,3]*, Jeff M.P. Holly[2]*.

Affiliations:

[1]Medical Research Council (MRC) Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom.

[2]IGFs & Metabolic Endocrinology Group, School of Translational Health Sciences, Learning & Research Building, Southmead Hospital, Bristol, United Kingdom.

[3]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom.

[4]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, 10461, USA.

[5]Institute for Community Medicine, University Medicine Greifswald, Greifswald, 17475, Germany.

[6]National Institute for Health Research (NIHR) Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and the University of Bristol.

[#]Members from the PRACTICAL Consortium (http://PRACTICAL.icr.ac.uk) are provided in the Supplementary Materials.

Correspondence to: Professor Nicholas J. Timpson, Address: MRC Integrative Epidemiology Unit, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, Tel: 0117 3310131.

Conflict of interest: The authors declare no potential conflicts of interest.

Abstract

Background:

Circulating insulin-like growth factor binding protein 3 (IGFBP-3) has been associated with PCa. Preclinical studies found that vitamin D regulates IGFBP-3 expression, although evidence from epidemiological studies are conflicting.

Methods:

Mendelian Randomisation analyses (MR) were conducted to reassess associations between IGFBP-3 and Prostate cancer (PCa) risk and advanced PCa using summary statistics from the PRACTICAL consortium (44,825cases;27,904controls). Observational and MR analyses were conducted to assess the relationship between inactive vitamin D (25(OH)D) and IGFBP-3 using data from the ProtecT study (1,366cases;1,071controls) and summary statistics from the CHARGE consortium (n=18,995).

Results:

The odds ratio (OR) for PCa per standard deviation (SD) unit increase in circulating IGFBP-3 was 1.14(95%CI:1.02-1.28). The OR for advanced PCa per SD unit increase in IGFBP-3 was 1.22(95%CI:1.07-1.40). Observationally, a SD increase in 25(OH)D was associated with a 0.1SD(95%CI:0.05, 0.14) increase in IGFBP-3. MR analyses found little evidence for a causal relationship between circulating 25(OH)D and IGFBP-3 in the circulation.

Conclusions:

This study provided confirmatory evidence that IGFBP-3 is a risk factor for PCa risk and progression. Observationally, there was evidence that 25(OH)D is associated with IGFBP-3, but MR analyses suggested that these findings were unlikely to be causal. Findings may be limited by the nature of instrumentation of 25(OH)D and IGFBP-3 and the utility of circulating measures. 25(OH)D appears unlikely to be causally related to IGFBP-3 in the circulation, however, our findings do not preclude causal associations at the tissue level.

Impact:

IGFBP-3 is a PCa risk factor but 25(OH)D are unlikely to be causally related to IGFBP-3 in the circulation.

Introduction

Insulin-like growth factor binding protein 3 (IGFBP-3) is the most abundant circulating IGFBP and modulates the bioactivity of IGFs[1]. Independent of IGFs, IGFBP-3 regulates cell proliferation, leading to increased interest in its role in carcinogenesis[2]. A meta-analysis found that circulating IGFBP-3 was associated with increased odds of Prostate cancer (PCa) (OR for highest versus lowest quintile: 1.25;95%CI:1.12,1.40)[3]. This is in line with results from a nested case-control study within the Prostate Testing for Cancer and Treatment (ProtecT) trial, which found evidence that circulating IGFBP-3 was associated with increased PCa odds (OR: 1.28 per SD increase in IGFBP-3; 95%CI:1.21,1.36)[4]. There is inconsistent evidence regarding the association between circulating IGFBP-3 and aggressive PCa. However, a Mendelian randomization (MR) study concluded that the IGFs may be associated with more aggressive PCa, in particular IGFBP-3[5].

Active vitamin D (1,25-hydroxyvitamin D;1,25$(OH)_2$D) up-regulates IGFBP-3 expression by binding to the vitamin D receptor which in turn binds to the vitamin D response element in the *IGFBP-3* promoter[6,7]. In the context of this relationship, there is motivation to assess the relationship between IGFBP-3 and the inactive form of vitamin D (25-hydroxyvitamin D;25(OH)D), as it acts as a proxy for 1,25$(OH)_2$D and is the most likely form of supplementation[8,9]. Evidence from observational studies examining the direct association between circulating 25(OH)D and 1,25$(OH)_2$D with PCa incidence and progression been inconclusive[10]. However, there is evidence that SNPs in vitamin D pathway are associated with higher PCa grade[11]. Despite this, a recent MR study using single nucleotide polymorphisms (SNPs) (representing lower 25(OH)D) failed to provide evidence of a causal association between 25(OH)D and risk of 7 types of cancers including PCa[12]. Currently, it is unlikely that 25(OH)D exerts large linear effects on PCa, but small or non-linear effects cannot be ruled out[13]. Indeed, preclinical studies have suggested that a possible mechanism for a vitamin D effect could be through the regulation of IGFBP-3[14-16].

Evidence from observational studies investigating the association between circulating 25(OH)D and IGFBP-3 has been conflicting. To compound this, observational studies can suffer from the effects of confounding, bias and reverse causation[17-20]. Specifically reverse causation could exist as IGFBP-3 can affect 25(OH)D through its effect on IGF-I and residual confounding might be present due to inadequate control for confounders such as BMI. Further examination of the relationship between

25(OH)D and IGFBP-3 is needed to clarify if vitamin D affects cancer via IGFBP-3.

Mendelian randomization (MR), which uses genetic variants as proxies of an exposure, is a causal analysis method used to investigate causality between exposures and health outcomes. Associations between SNPs and outcomes can provide evidence of causation since they are not subjected to bias (i.e.,reverse causation and confounding) found in observational studies[21]. Recently, two meta-analyses of Genome-Wide Association Studies (GWAS) identified SNPs associated with circulating IGFBP-3[22] and 25(OH)D[23], respectively. The SNPs associated with circulating IGFBP-3 and 25(OH)D provide a framework to assess the consequence of lifelong 25(OH)D on IGFBP-3 independent of other factors.

In this study, we aimed to re-assess the causal relationship between IGFBP-3 and PCa. We also aimed to undertake both observational and one-sample MR analyses within ProtecT to assess the observational and causal relationship between 25(OH)D and IGFBP-3 and to compare these in order to assess the impact of potential bias, confounding or reverse causality. We also aimed to investigate whether there is a causal effect of 25(OH)D on IGFBP-3 using two-sample MR in independent data from the IGF working group of the CHARGE consortium.

Methods and Materials

Study Design

This investigation had three components (**Figure 1**): 1) two-sample MR analyses were conducted to reassess the causal relationship between IGFBP-3 and PCa risk and advanced PCa (Gleason≥8, death from PCa, PSA>100 or metastasis) using summary statistics from the PRACTICAL consortium; 2) linear regression analyses were conducted to investigate the observational relationship between circulating 25(OH)D and IGFBP-3 using individual level data from ProtecT; and 3) one-sample MR analyses using individual level data from ProtecT and two-sample MR analyses[24,25] using summary statistics from the IGFBP-3 GWAS were conducted to investigate the causal relationship between 25(OH)D and IGFBP-3.

Study populations and data sources for observational and MR analyses

Individual level analyses (observational and one-sample MR) were examined in a nested case-control

study within the ProtecT trial[26,27] (details of the cohort are described in the **Supplementary methods**). For the observational analyses, 1,446 cases and 1,449 controls with 25(OH)D and IGFBP-3 measured at diagnosis and before treatment were included. Circulating 25(OH)D in blood plasma was measured using tandem mass spectrometry, as described previously[28]. Circulating $1,25(OH)_2D$ were measured in blood plasma using immunoassay, as previously described[29]. Circulating IGFBP-3 in blood serum was measured using in-house radioimmunoassay, as previously described[4] (see **Supplementary methods**). Seasonality in 25(OH)D and $1,25(OH)_2D$ levels were adjusted using the cosinor method as previously described[30]. This method was chosen instead of adjusting models for season of blood draw, as it is a more accurate way to assess an individual's vitamin D status[31]. Details of how potential confounders (BMI, exercise, smoking, alcohol consumption, family history of PCa, history of benign prostatic hyperplasia, diabetes, social class and ethnicity) were measured are provided in the **Supplementary methods.** Genome-wide genotyping of participants was carried out using the Illumina Human660W array (see **Supplementary methods**). Altogether, 674 cases and 410 controls with genotype data and IGFBP-3 measured were included in the one-sample MR analyses.

A previous MR study[5] had investigated the association between the IGF axis with risk of PCa by identifying SNPs associated with the IGFs from a GWAS for IGFBP-3[32] (n=10,018) and candidate gene studies and using summary statistics from a PCa risk GWAS (n=22,898 cases and 23,054 controls) genotyped using the iCOGS array. A new GWAS for risk of PCa[33] (n=44,825 cases and 27,904 controls) genotyped using Oncoarray and an up-dated GWAS for IGFBP-3[22] (n=18,995) were recently published. We up-dated the previous MR analysis for IGFBP-3 and risk of PCa using the latest PCa GWAS meta-analysis[33] and the latest IGFBP-3 GWAS meta-analysis[22] (see **Supplementary methods**). For our MR analyses investigating the causal relationship between 25(OH)D and IGFBP-3, we used the following summary statistics: 1) IGFBP-3 GWAS[22] conducted by the IGF working group of the CHARGE consortium (n=18,995); and 2) 25(OH)D GWAS[34] (n=79,366) conducted by the SUNLIGHT consortium (see **Supplementary methods**).

Identification of genetic instruments for MR analyses

Instruments for IGFBP-3

The previous MR study[5] investigating the causal relationship between IGFs and PCa used individual SNPs identified either by the discovery IGF GWAS[32] or by candidate gene studies. For our MR

analyses, we constructed a new instrument for IGFBP-3 by selecting four independent IGFBP-3 associated SNPs (rs11977526, rs700753, rs1065656, rs4234798) identified by the up-dated IGFBP-3 GWAS[22]. rs700753 was not present in the 1000 genomes imputation reference used in the PCa GWAS. We identified a proxy SNP, rs700752, which was in high linkage disequilibrium (r^2>0.8) using SNAP (https://www.broadinstitute.org/mpg/snap/).

Instruments for 25(OH)D

To-date, GWAS studies[23,35,36] have focused on identifying SNPs associated with 25(OH)D as it is more stable than 1,25$(OH)_2$D in the circulation. The latest 25(OH)D GWAS meta-analysis[34] conducted in Europeans identified 2 novel loci (*SEC23A* (rs8018720) and *AMDHD1* (rs10745742)) with a genome-wide significant association with 25(OH)D and confirmed 4 previously identified loci[23] (CYP2R1(rs10741657);DHCR7(rs12785878);GC(rs3755967);and CYP24A1(rs17216707)) located in or near genes involved in 25(OH)D synthesis and metabolism. We constructed an instrument for 25(OH)D (allele score) by using the six 25(OH)D-associated SNPs. As sensitivity analyses, we constructed a synthesis (rs12785878 and rs10741657) and metabolism (rs3755967 and rs17216707) score using SNPs in genes involved in 25(OH)D synthesis and metabolism.

Statistical analyses

Causal association between IGFBP-3 with overall cancer risk and advanced PCa

SNP-exposure (IGFBP-3) and SNP-outcome (PCa) estimates for the four IGFBP-3 associated SNPs were combined using inverse-variance weighted (IVW) and maximum likelihood method to provide a weighted average of the causal estimates[37]. The previous MR study found that SNPs associated with IGFBP-3 have pleiotropic effects on other biomarkers of the IGF pathway[5]. The latest IGF GWAS found that two IGFBP-3 associated SNPs (rs700753 and rs1065656) were also associated with IGF-I at genome-wide significant levels[22]. To account for the effect of the IGFBP-3 associated SNPs on IGF-I, we employed Multivariable MR, a new MR method that estimates causal effects using multiple SNPs associated with multiple exposures simultaneously[38,39]. Multivariable MR was conducted by regression of the SNP-PCa estimates on SNP-IGFBP-3 and SNP-IGF-I estimates in a multivariable weighted regression model. As the IGF GWAS did not analyse other IGFs, the multivariable MR method could not take into account the association between SNPs and other biomarkers of the IGF pathway.

Observational relationship between circulating 25(OH)D and IGFBP-3 in ProtecT

Observational associations between 25(OH)D and IGFBP-3 were assessed using linear regression for cases and controls combined into one cohort or stratified by case-control status. Additional analyses were adjusted for potential confounders (age, centre, bmi, smoking and diabetes status). Associations of 25(OH)D and IGFBP-3 with potential confounders (listed above) were estimated using linear regression. Potential non-linear effects of 25(OH)D and IGFBP-3 were tested by carrying out linear regression with quadratic terms. Details of the analysis investigating the observational relationship between circulating $1,25(OH)_2D$ and IGFBP-3 in ProtecT are provided in the **Supplementary methods.**

Causal associations between 25(OH)D and IGFBP-3: one-sample MR in ProtecT

A weighted genetic risk score (GRS) for 25(OH)D was generated by taking the sum of genotypes (coded 0,1,2) multiplied by the strength of the effect of each SNP on 25(OH)D recorded by the 25(OH)D GWAS[23]. Two-stage least squares analyses using the weighted GRS was used to obtain estimates of the associations between season-adjusted 25(OH)D and IGFBP-3[40,41]. The properties of the 25(OH)D-associated SNPs as instruments were assessed in ProtecT by examination of: i) first stage F-statistic (measure of the strength of the association between the instrument and exposure) and ii) associations of the SNPs with potential confounders. The first-stage F-statistic was obtained from the regression of 25(OH)D on the 25(OH)D genetic instrument (first stage of the two-stage least squares MR analysis). Sensitivity analyses were also performed by combining the estimates from SNPs involved in the synthesis and metabolism of vitamin D separately. We compared the instrumental variable estimates with those from observational analyses using the Durbin form of the Durbin-Wu-Hausman statistic[42].

Causal associations between 25(OH)D and IGFBP-3: two-sample MR

The SNP-exposure (25(OH)D) and SNP-outcome (IGFBP-3) estimates for the six 25(OH)D associated SNPs were combined using the IVW method[37]. The IVW method (similar method to that used in the two-sample MR analysis of IGFBP-3 and PCa) is equivalent to the two-stage least square analyses using data from ProtecT[43]. Sensitivity analyses were performed by combining the estimates from SNPs involved in the synthesis and metabolism pathways separately. As only a small number of SNPs were included in our MR analyses, methods to test for pleiotropy such as MR-Egger

regression[43], weighted median[44] and mode method[45] were not conducted as these lack power with small number of SNPs. A leave-one-out analysis, which repeats the MR analysis by leaving out each SNP in turn, was applied to assess whether any single SNP was driving the causal estimate. In the reverse direction, we investigated the causal effect of IGFBP-3 on 25(OH)D using the two-sample MR analyses (see **Supplementary methods**).

Power

Power for one-sample MR analyses to detect the causal effect of 25(OH)D on IGFBP-3 was estimated using the online tool (http://glimmer.rstudio.com/kn3in/mRnd/)[46], specifying alpha=0.05, R^2=0.05[35] and using the observational estimates from the ProtecT sample set. Given a sample size of 1134, in one-sample MR, a causal effect of 0.39 SD units per SD unit change in 25(OH)D was required to yield 80% power. For two-sample MR analysis of the same relationship, power to detect the causal effect of 25(OH)D on IGFBP-3 was estimated using the "pwr.r.test" function (implemented in the pwr package in R) specifying alpha=0.05. The correlation between 25(OH)D and IGFBP-3 was obtained by taking the square-root of the variance in IGFBP-3 explained by the 25(OH)D associated SNPs. Given a sample size of 18,995 in the two-sample MR analysis, a causal effect needed to yield 80% power is 0.09 SD units. This method is not dissimilar to that suggested elsewhere noting that power in a two-sample procedure is limited by the direct SNP-outcome association[47]. All analyses were performed in R (version 3.0.1), Stata version 14 (Stata Corp) and using MR-Base (www.mrbase.org).

Results

Association between IGFBP-3 with overall cancer risk and advanced PCa

In agreement with the previous MR of IGFBP-3 and PCa[5], two-sample MR analyses using the latest PCa GWAS found that the OR for PCa per SD unit increase in circulating IGFBP-3 was 1.12 (95%CI:0.91-1.36;*p*=0.18) from IVW analyses. The multivariable estimate, which took into account the effect of the SNPs on IGF-I, was of similar magnitude to the IVW analyses and gave stronger evidence with an OR of 1.14 (95%CI:1.02-1.28;*p*=0.02) (**Supplementary Table 1)**. The OR for advanced PCa per SD unit increase in IGFBP-3 was 1.16 (95%CI:0.87-1.55;*p*=0.19) from IVW analyses. The multivariable estimate was of similar magnitude to the IVW analyses, but gave stronger evidence with an OR of 1.22 (95%CI:1.07-1.40;*p*=0.004) (**Supplementary Table 1)**.

Observational association between circulating 25(OH)D and IGFBP-3 in ProtecT

Mean circulating IGFBP-3 levels were higher in the cases compared to the controls in ProtecT (4634.77ng/ml versus 4502.19ng/ml, *p* for difference=0.002). Mean season-adjusted circulating 25(OH)D levels did not differ between cases and controls (22.75 ng/ml versus 22.71 ng/ml, respectively, *p* for difference=0.87). Mean season-adjusted $1,25(OH)_2D$ levels did not differ between cases and controls (40.71pg/ml versus 42.06pg/ml, respectively, *p* for difference=0.12) (**Table 1**). For cases and controls combined, circulating 25(OH)D and season-adjusted 25(OH)D were positively correlated with $1,25(OH)_2D$ levels (r=0.24;*p*<0.001 and r=0.22;*p*<0.001, respectively) (**Supplementary Table 2**).

For cases and controls combined, an SD unit increase (SD=7.8ng/ml) in circulating 25(OH)D was associated with a 0.09 SD (95%CI:0.05,0.13;*p*<0.001) increase in circulating IGFBP-3 levels. A 0.09 SD unit increase in IGFBP-3 is equivalent to 93.3ng/ml. There was no strong evidence for difference between cases and controls (**Table 2**). We found evidence that circulating 25(OH)D and IGFBP-3 are associated with potential confounding factors including age, BMI and diabetes status (**Supplementary Table 3**). After adjustment for confounders, the effect (beta: 0.09; 95% CI: 0.04, 0.14; *p*<0.001) was similar in magnitude to the unadjusted analysis. When stratified by case-control status, the precision of the estimates was reduced due to smaller sample sizes (**Table 2**). We found little evidence for a non-linear effect of vitamin D on IGFBP-3 (*p*=0.43).

For cases and controls combined, an SD unit increase in circulating $1,25(OH)_2D$ was associated with a 0.09 SD (95%CI:0.04,0.14;*p*=0.001) increase in IGFBP-3 (**Table 2**). We found evidence for association between $1,25(OH)_2D$ and potential confounders including BMI, family history of PCa, smoking and diabetes status (**Supplementary Table 3**). After adjustment for potential confounders, the effect (beta: 0.09; 95%CI: 0.03, 0.15; *p*=0.003) was similar in magnitude to the unadjusted analysis. When stratified by case-control status, the precision of the estimates was reduced due to smaller sample sizes (**Table 2**).

Validation of the instruments for circulating 25(OH)D in ProtecT

A MR assumption is that there is a reliable association between the genetic instrument and the exposure[48]. We validated this using data from ProtecT in which 1,416 individuals had both genotype

data and 25(OH)D measured. In cases and controls combined, per unit increase in the allele score was associated with a 0.11 SD unit increase in 25(OH)D. There was no strong evidence for difference between cases and controls. For an individual with the maximum number of 25(OH)D increasing alleles (12), this corresponds to a 1.32 SD unit difference (on average) in 25(OH)D when compared to an individual with the minimum (0) number of 25(OH)D increasing alleles. This is higher in magnitude and is comparable to that of vitamin D supplementation which increases 25(OH)D by 0.45 SD units[35]. In cases and controls combined, the metabolism score was associated with increased 25(OH)D (p<0.001). For the metabolism score, there was no strong evidence for difference between cases and controls. In cases and controls combined, the synthesis score was found to be associated with increased 25(OH)D; however, the precision of the estimates was reduced when stratified by case-control status due to reduced sample size (**Supplementary Table 4**).

Using the allele score, each additional 25(OH)D-related allele was associated with a 0.05 pg/ml increase in $1,25(OH)_2D$ in cases and controls combined. There was no strong evidence for difference between cases and controls. In cases and controls combined, the metabolism score was also found to be associated with increased $1,25(OH)_2D$ (p=0.005); however, the precision of the estimates was reduced when stratified by case-control status due to reduced sample size (**Supplementary Table 4).**

Another MR assumption is that the genetic instrument should not be associated with confounders (measured or unmeasured) of the exposure-outcome relationship[48]. There was little evidence that the allele, metabolism or synthesis score) were associated with potential confounders **(Supplementary Table 5**).

Causal associations between 25(OH)D and IGFBP-3: one-sample MR in ProtecT

The causal effect of 25(OH)D on IGFBP-3 was estimated using one-sample MR (**Figure 2).** Using the allele score as an instrument for 25(OH)D, for cases and controls combined, the instrumental variable (IV) estimate of the causal relationship between a one SD increase in 25(OH)D and IGFBP-3 for cases and controls combined was -0.32 SD units (95%CI:-0.64,-0.01;p=0.04). When stratified by case-control status, the IV estimate for the controls was -0.10 (95%CI:-0.62,0.41;p=0.69) and the IV estimate for the cases was -0.30 (95%CI:-0.79,-0.02; p=0.04). Sensitivity analyses based on the use of individual SNPs, synthesis and metabolism score showed little evidence that 25(OH)D was causally

associated with IGFBP-3 (**Table 3**;**Figure 3**). The first-stage F-statistic using the allele score and metabolism score was 46.76 and 41.12, respectively. There was evidence of a departure of instrumental variable-derived estimates from those derived from observational analyses (Durbin-Wu-Hausman test *p*=0.007) (**Table 3**).

Causal associations between 25(OH)D and IGFBP-3: two-sample MR

Using the six 25(OH)D-associated SNPs for the two-sample MR analysis, the IVW estimate between a one log-unit increase in 25(OH)D and IGFBP-3 was 0.11 SD units (95%CI:-0.10,0.31;*p*=0.32) (**Figure 3**). In leave-one-out analyses, sequentially omitting each of the six 25(OH)D associated SNPs provided similar causal estimates to IVW analysis (**Supplementary Table 6**).

Sensitivity analyses using SNPs involved in 25(OH)D synthesis found that the IVW estimate between a one log-unit increase in 25(OH)D and IGFBP-3 was 0.15 SD units (95%CI:-0.30,0.61;*p*=0.51). Using SNPs involved in 25(OH)D metabolism, the IVW estimate between a one log-unit increase in 25(OH)D and IGFBP-3 was 0.07 SD units (95%CI:-0.17,0.31;*p*=0.59) (**Figure 3**).

In the reverse direction, we investigated the causal effect of IGFBP-3 on 25(OH)D using the two-sample MR. The IVW estimate between a one SD increase in IGFBP-3 and 25(OH)D was 0.01 log-units (95%CI:-0.003,0.03;*p*=0.10) (**Supplementary Table 7**).

Discussion

This study supported results from previous studies that the IGF pathway is associated with increased odds of PCa and of advanced PCa - in particular IGFBP-3[4,5]. Existing preclinical research has implicated $1,25(OH)_2D$ in the regulation of IGFBP-3 expression and in this context we set out to assess the causal relationship between the closest modifiable proxy for this (25(OH)D) and IGFBP-3. Using data from ProtecT, this study found observational evidence that 25(OH)D and $1,25(OH)_2D$ are associated with IGFBP-3 in the circulation. However, MR analyses using data from ProtecT were not suggestive of a causal effect for 25(OH)D. Using summary statistics from the new IGFBP-3 GWAS, two-sample MR analyses found little evidence of a causal relationship between circulating 25(OH)D and IGFBP-3.

The relationship of vitamin D and IGFBP-3 is complex due to the co-existing endocrine and autocrine/paracrine interactions[6]. At the autocrine/paracrine level, several preclinical studies have demonstrated that 1,25$(OH)_2$D regulates IGFBP-3 expression[6,7] and that the growth inhibitory effects of 1,25$(OH)_2$D on PCa cells is altered by IGFBP-3[14-16]. At the endocrine level, evidence from observational studies examining whether 25(OH)D or 1,25$(OH)_2$D influence IGFBP-3 in the circulation have been conflicting, but inferences of causality are limited with observational studies due to effects of confounding and reverse causation[17-20]. Our MR results suggesting that 25(OH)D does not have a causal effect on IGFBP-3 levels in the circulation are in agreement with the results from a randomized phase II trial of men with localised PCa and high grade prostatic intraepithelial neoplasia which found that supplementation with the vitamin D analog,1-alpha-hydroxyvitamin D2, did not have an effect on circulating IGFBP-3 levels[49]. Another vitamin D supplementation trial conducted among women with high risk of breast cancer also found little evidence that circulating IGFBP-3 levels changed after supplementation with 25(OH)D[50]. This does not preclude that increasing the availability of 25(OH)D with supplementation may lead to increased production of 1,25$(OH)_2$D within tissues that could then stimulate the local production of IGFBP-3 that is not detected in the circulation.

Evidence from this study corroborates that the IGF pathway is a risk factor for PCa risk and progression[3-5], but challenges assumptions about the possible route from supplement-derived precursor vitamin D, through IGFBP-3 to cancer. This work sits in the broader context of a prospective nested case-control study which found evidence for interactions between plasma 25(OH)D and plasma IGF-I/IGFBP-3 ratio with regard to colorectal cancer risk[20]. Another study of premenopausal women also reported that vitamin D supplementation was negatively associated with mammographic breast density, and that the association was stronger in those with high IGF-I and IGFBP-3 levels[18].

There could be several reasons why our MR findings differ from results from the observational studies. Firstly, the point estimates are close to the null for our 2 sample MR analyses. It is possible that the true effect of 25(OH)D on IGFBP-3 could be small and undetectable using a causal analysis framework. Secondly, the genetic instruments for 25(OH)D may not be acting as a true proxy for 1,25$(OH)_2$D. Our analyses using the allele and metabolism score found that each additional 25(OH)D-related allele was associated with increased 1,25$(OH)_2$D levels for controls and cases combined (**Supplementary Table 4**). However, the F-statistic for instrumental analyses using allele, metabolism

and synthesis scores was lower for $1,25(OH)_2D$ compared to 25(OH)D (**Table 3**), indicating that they are weaker instruments for $1,25(OH)_2D$. Our observational analyses in ProtecT found that $1,25(OH)_2D$ is associated with IGFBP-3 in the circulation; however, it is still unclear if a causal relationship exists between circulating $1,25(OH)_2D$ and IGFBP-3. Thirdly, given that vitamin D activating enzymes are present in many tissues and could influence $1,25(OH)_2D$ concentrations within tissues, including prostate tissues[51,52], circulating 25(OH)D and $1,25(OH)_2D$ may not reflect the true bioavailability of 25(OH)D and $1,25(OH)_2D$ within tissues. Therefore, the endocrine relationship between 25(OH)D and IGFBP-3 may not be reflective of the autocrine/paracrine relationship.

Our study has several limitations. The SNPs associated with IGFBP-3 have been shown to be associated with more than one part of the IGF pathway. Although the multivariable MR method enables the determination of the causal effect of IGFBP-3 on PCa independent of IGF-I, we cannot rule out associations with other members of the IGF pathway[5]. As only a small number of SNPs were included in our MR analyses, methods to test for pleiotropy such as the MR-Egger[43], weighted median[44] and mode[45] method were not conducted as these lack power with small number of SNPs. As the SNPs associated with 25(OH)D only explained a small proportion of the variability in circulating 25(OH)D, large sample sizes are required to detect expected influences on IGFBP-3 levels and to be able to provide precise estimates for this effect. Although our point estimates are close to the null for the two-sample MR analyses, we cannot rule out the possibility that 25(OH)D may have small effects on IGFBP-3. Lastly, our study population was limited to Europeans. Although population homogeneity eliminates population admixture as a potential confounder in our analyses, as 25(OH)D levels vary with sun exposure and skin colour, the findings drawn from this study might not be applicable to other ethnic groups or individuals in different geographical locations.

Conclusions

This study confirmed results from previous studies that members of the IGF pathway, in particular IGFBP-3, has a causal effect on PCa and advanced PCa. Observationally, there is evidence that circulating 25(OH)D is positively associated with circulating IGFBP-3, but MR analyses indicated these findings were unlikely to be causal. Findings here are limited by the nature of instrumentation of both 25(OH)D and IGFBP-3 and the utility of circulating measures, but are important as they suggest that circulating 25(OH)D are unlikely to be causally related to circulating IGFBP-3, although vitamin D

supplementation could affect local tissue levels.

Acknowledgements

This work was supported by CRUK [The Integrative Cancer Epidemiology Programme, C18281/A19169]. VYT, TD, CB, NJT and RMM work in the Medical Research Council Integrative Epidemiology Unit at the University of Bristol which is supported by the Medical Research Council and the University of Bristol [grant number MC_UU_12013/3]. NJT is a Wellcome Trust Investigator (202802/Z/16/Z) and works within the University of Bristol NIHR Biomedical Research Centre (BRC). TD is funded by The Wellcome Trust [grant number 201237/Z/16/Z], KMB is funded by Cancer Research UK (C11046/A10052) and by the National Institute for Health Research (NIHR) Bristol Biomedical Research Centre based at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. QQ is supported by a Scientist Development Award (K01HL129892) from the NHLBI .The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

References

1. Johnson MA, Firth SM. IGFBP-3: a cell fate pivot in cancer and disease. *Growth hormone & IGF research : official journal of the Growth Hormone Research Society and the International IGF Research Society* 2014;24(5):164-73. doi: 10.1016/j.ghir.2014.04.007 [published Online First: 2014/06/24]
2. Firth SM, Baxter RC. Cellular actions of the insulin-like growth factor binding proteins. *Endocrine reviews* 2002;23(6):824-54. doi: 10.1210/er.2001-0033 [published Online First: 2002/12/06]
3. Travis RC, Appleby PN, Martin RM, Holly JM, Albanes D, Black A, et al. A Meta-analysis of Individual Participant Data Reveals an Association between Circulating Levels of IGF-I and Prostate Cancer Risk. *Cancer research* 2016;76(8):2288-300. doi: 10.1158/0008-5472.CAN-15-1551 [published Online First: 2016/02/28]
4. Rowlands MA, Holly JM, Gunnell D, Donovan J, Lane JA, Hamdy F, et al. Circulating insulin-like growth factors and IGF-binding proteins in PSA-detected prostate cancer: the large case-control study ProtecT. *Cancer research* 2012;72(2):503-15. doi: 10.1158/0008-5472.CAN-11-1601 [published Online First: 2011/11/23]
5. Bonilla C, Lewis SJ, Rowlands MA, Gaunt TR, Davey Smith G, Gunnell D, et al. Assessing the role of insulin-like growth factors and binding proteins in prostate cancer using Mendelian randomization: Genetic variants as instruments for circulating levels. *International journal of cancer* 2016;139(7):1520-33. doi: 10.1002/ijc.30206 [published Online First: 2016/05/27]
6. Ameri P, Giusti A, Boschetti M, Murialdo G, Minuto F, Ferone D. Interactions between vitamin D and IGF-I: from physiology to clinical practice. *Clinical endocrinology* 2013;79(4):457-63. doi: 10.1111/cen.12268 [published Online First: 2013/06/25]
7. Peng L, Malloy PJ, Feldman D. Identification of a functional vitamin D response element in the human insulin-like growth factor binding protein-3 promoter. *Molecular endocrinology* 2004;18(5):1109-19. doi: 10.1210/me.2003-0344 [published Online First: 2004/02/14]
8. Dubbelman R, Jonxis JH, Muskiet FA, Saleh AE. Age-dependent vitamin D status and vertebral condition of white women living in Curacao (The Netherlands Antilles) as compared with their counterparts in The Netherlands. *The American journal of clinical nutrition* 1993;58(1):106-9. [published Online First: 1993/07/01]
9. Lips P, Wiersinga A, van Ginkel FC, Jongen MJ, Netelenbos JC, Hackeng WH, et al. The effect of vitamin D supplementation on vitamin D status and parathyroid function in elderly subjects. *The Journal of clinical endocrinology and metabolism* 1988;67(4):644-50. doi: 10.1210/jcem-67-4-644 [published Online First: 1988/10/01]
10. Jacobs ET, Kohler LN, Kunihiro AG, Jurutka PW. Vitamin D and Colorectal, Breast, and Prostate Cancers: A Review of the Epidemiological Evidence. *Journal of Cancer* 2016;7(3):232-40. doi: 10.7150/jca.13403 [published Online First: 2016/02/27]
11. Gilbert R, Bonilla C, Metcalfe C, Lewis S, Evans DM, Fraser WD, et al. Associations of vitamin D pathway genes with circulating 25-hydroxyvitamin-D, 1,25-dihydroxyvitamin-D, and prostate cancer: a nested case-control study. *Cancer causes & control : CCC* 2015;26(2):205-18. doi: 10.1007/s10552-014-0500-5 [published Online First: 2014/12/10]
12. Dimitrakopoulou VI, Tsilidis KK, Haycock PC, Dimou NL, Al-Dabhani K, Martin RM, et al. Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study. *BMJ* 2017;359:j4761. doi: 10.1136/bmj.j4761

13. Manousaki D, Richards JB. Low vitamin D levels as a risk factor for cancer. *BMJ* 2017;359:j4952. doi: 10.1136/bmj.j4952
14. Krishnan AV, Shinghal R, Raghavachari N, Brooks JD, Peehl DM, Feldman D. Analysis of vitamin D-regulated gene expression in LNCaP human prostate cancer cells using cDNA microarrays. *The Prostate* 2004;59(3):243-51. doi: 10.1002/pros.20006 [published Online First: 2004/03/26]
15. Boyle BJ, Zhao XY, Cohen P, Feldman D. Insulin-like growth factor binding protein-3 mediates 1 alpha,25-dihydroxyvitamin d(3) growth inhibition in the LNCaP prostate cancer cell line through p21/WAF1. *The Journal of urology* 2001;165(4):1319-24. [published Online First: 2001/03/21]
16. Stewart LV, Weigel NL. Role of insulin-like growth factor binding proteins in 1alpha,25-dihydroxyvitamin D(3)-induced growth inhibition of human prostate cancer cells. *The Prostate* 2005;64(1):9-19. doi: 10.1002/pros.20212 [published Online First: 2005/01/15]
17. DeLellis K, Rinaldi S, Kaaks RJ, Kolonel LN, Henderson B, Le Marchand L. Dietary and lifestyle correlates of plasma insulin-like growth factor-I (IGF-I) and IGF binding protein-3 (IGFBP-3): the multiethnic cohort. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2004;13(9):1444-51. [published Online First: 2004/09/03]
18. Diorio C, Berube S, Byrne C, Masse B, Hebert-Croteau N, Yaffe M, et al. Influence of insulin-like growth factors on the strength of the relation of vitamin D and calcium intakes to mammographic breast density. *Cancer research* 2006;66(1):588-97. doi: 10.1158/0008-5472.CAN-05-1959 [published Online First: 2006/01/07]
19. Holmes MD, Pollak MN, Willett WC, Hankinson SE. Dietary correlates of plasma insulin-like growth factor I and insulin-like growth factor binding protein 3 concentrations. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2002;11(9):852-61. [published Online First: 2002/09/12]
20. Wu K, Feskanich D, Fuchs CS, Chan AT, Willett WC, Hollis BW, et al. Interactions between plasma levels of 25-hydroxyvitamin D, insulin-like growth factor (IGF)-1 and C-peptide with risk of colorectal cancer. *PloS one* 2011;6(12):e28520. doi: 10.1371/journal.pone.0028520 [published Online First: 2012/01/05]
21. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human molecular genetics* 2014;23(R1):R89-98. doi: 10.1093/hmg/ddu328 [published Online First: 2014/07/30]
22. Teumer A, Qi Q, Nethander M, Aschard H, Bandinelli S, Beekman M, et al. Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. *Aging cell* 2016;15(5):811-24. doi: 10.1111/acel.12490 [published Online First: 2016/06/23]
23. Manousaki D, Dudding T, Haworth S, Hsu Y, Liu C, Medina-Gómez C, et al. Low frequency coding variation in CYP2R1 has large effects on vitamin D level and risk of multiple sclerosis *Manuscript submitted for publication* 2017
24. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Statistics in medicine* 2008;27(8):1133-63. doi: 10.1002/sim.3034 [published Online First: 2007/09/22]

25. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG, Consortium E-I. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *European journal of epidemiology* 2015;30(7):543-52. doi: 10.1007/s10654-015-0011-z [published Online First: 2015/03/17]
26. Lane A, Metcalfe C, Young GJ, Peters TJ, Blazeby J, Avery KN, et al. Patient-reported outcomes in the ProtecT randomized trial of clinically localized prostate cancer treatments: study design, and baseline urinary, bowel and sexual function and quality of life. *BJU international* 2016;118(6):869-79. doi: 10.1111/bju.13582 [published Online First: 2016/07/15]
27. Lane JA, Hamdy FC, Martin RM, Turner EL, Neal DE, Donovan JL. Latest results from the UK trials evaluating prostate cancer screening and treatment: the CAP and ProtecT studies. *European journal of cancer* 2010;46(17):3095-101. doi: 10.1016/j.ejca.2010.09.016 [published Online First: 2010/11/05]
28. Gilbert R, Metcalfe C, Fraser WD, Donovan J, Hamdy F, Neal DE, et al. Associations of circulating 25-hydroxyvitamin D with prostate cancer diagnosis, stage and grade. *International journal of cancer* 2012;131(5):1187-96. doi: 10.1002/ijc.27327 [published Online First: 2011/10/29]
29. Gilbert R, Metcalfe C, Fraser WD, Donovan J, Hamdy F, Neal DE, et al. Associations of circulating retinol, vitamin E, and 1,25-dihydroxyvitamin D with prostate cancer diagnosis, stage, and grade. *Cancer causes & control : CCC* 2012;23(11):1865-73. doi: 10.1007/s10552-012-0052-5 [published Online First: 2012/08/29]
30. Lawlor DA, Wills AK, Fraser A, Sayers A, Fraser WD, Tobias JH. Association of maternal vitamin D status during pregnancy with bone-mineral content in offspring: a prospective cohort study. *Lancet* 2013;381(9884):2176-83. doi: 10.1016/S0140-6736(12)62203-X [published Online First: 2013/03/23]
31. Degerud E, Hoff R, Nygard O, Strand E, Nilsen DW, Nordrehaug JE, et al. Cosinor modelling of seasonal variation in 25-hydroxyvitamin D concentrations in cardiovascular patients in Norway. *European journal of clinical nutrition* 2016;70(4):517-22. doi: 10.1038/ejcn.2015.200 [published Online First: 2015/11/26]
32. Kaplan RC, Petersen AK, Chen MH, Teumer A, Glazer NL, Doring A, et al. A genome-wide association study identifies novel loci associated with circulating IGF-I and IGFBP-3. *Human molecular genetics* 2011;20(6):1241-51. doi: 10.1093/hmg/ddq560
33. Schumacher FR AAOA, Berndt SI, et al. Prostate cancer meta-analysis of more than 140,000 men identifies 63 novel prostate cancer susceptibility loci. *Nature Genetics* 2017;In Press
34. Jiang X, O'Reilly PF, Aschard H, Hsu YH, Richards JB, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat Commun* 2018;9(1):260. doi: 10.1038/s41467-017-02662-2
35. Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet* 2010;376(9736):180-8. doi: 10.1016/S0140-6736(10)60588-0 [published Online First: 2010/06/15]
36. Ahn J, Yu K, Stolzenberg-Solomon R, Simon KC, McCullough ML, Gallicchio L, et al. Genome-wide association study of circulating vitamin D levels. *Human molecular genetics* 2010;19(13):2739-45. doi: 10.1093/hmg/ddq155 [published Online First: 2010/04/27]

37. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology* 2013;37(7):658-65. doi: 10.1002/gepi.21758 [published Online First: 2013/10/12]
38. Burgess S, Dudbridge F, Thompson SG. Re: "Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects". *American journal of epidemiology* 2015;181(4):290-1. doi: 10.1093/aje/kwv017 [published Online First: 2015/02/11]
39. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *American journal of epidemiology* 2015;181(4):251-60. doi: 10.1093/aje/kwu283 [published Online First: 2015/01/30]
40. Greenland S. An introduction to instrumental variables for epidemiologists. *International journal of epidemiology* 2000;29(4):722-9. [published Online First: 2000/08/03]
41. Timpson NJ, Lawlor DA, Harbord RM, Gaunt TR, Day IN, Palmer LJ, et al. C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. *Lancet* 2005;366(9501):1954-9. doi: 10.1016/S0140-6736(05)67786-0 [published Online First: 2005/12/06]
42. Baum C SM, Stillman S. IVENDOG: Stata module to calculate Durbin-Wu-Hausman endogeneity test after ivreg. Statistical Software Components S494401: Boston: Boston College Department of Economics.; 2007 [
43. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *International journal of epidemiology* 2015;44(2):512-25. doi: 10.1093/ije/dyv080 [published Online First: 2015/06/08]
44. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic epidemiology* 2016;40(4):304-14. doi: 10.1002/gepi.21965
45. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *International journal of epidemiology* 2017;46(6):1985-98. doi: 10.1093/ije/dyx102
46. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *International journal of epidemiology* 2013;42(5):1497-501. doi: 10.1093/ije/dyt179 [published Online First: 2013/10/26]
47. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *American journal of epidemiology* 2013;178(7):1177-84. doi: 10.1093/aje/kwt084 [published Online First: 2013/07/19]
48. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *The American journal of clinical nutrition* 2016;103(4):965-78. doi: 10.3945/ajcn.115.118216 [published Online First: 2016/03/11]
49. Gee J, Bailey H, Kim K, Kolesar J, Havighurst T, Tutsch KD, et al. Phase II open label, multi-center clinical trial of modulation of intermediate endpoint biomarkers by 1alpha-hydroxyvitamin D2 in patients with clinically localized prostate cancer and high grade pin. *The Prostate* 2013;73(9):970-8. doi: 10.1002/pros.22644 [published Online First: 2013/01/22]

50. Crew KD, Xiao T, Thomas PS, Terry MB, Maurer M, Kalinsky K, et al. Safety, Feasibility, and Biomarker Effects of High-Dose Vitamin D Supplementation Among Women at High Risk for Breast Cancer. *International journal of food science, nutrition and dietetics* 2015;2015(Suppl 1):1-16. doi: 10.19070/2326-3350-SI01001 [published Online First: 2015/01/01]
51. Chen TC, Sakaki T, Yamamoto K, Kittaka A. The roles of cytochrome P450 enzymes in prostate cancer development and treatment. *Anticancer research* 2012;32(1):291-8. [published Online First: 2012/01/04]
52. Blomberg Jensen M, Nielsen JE, Jorgensen A, Rajpert-De Meyts E, Kristensen DM, Jorgensen N, et al. Vitamin D receptor and vitamin D metabolizing enzymes are expressed in the human male reproductive tract. *Human reproduction* 2010;25(5):1303-11. doi: 10.1093/humrep/deq024 [published Online First: 2010/02/23]

Table 1. Baseline characteristics of participants in the ProtecT study

Continuous variables	Controls (N=1449)		Cases (N=1446)	
	N	**Mean (SD or IQR)**	**N**	**Mean (SD or IQR)**
IGFBP-3 (ng/ml)	1071	4502.19 (1026.06)	1366	4634.77 (1041.58)
25(OH)D (ng/ml)	1449	22.93 (8.36)	1446	23.05 (8.71)
season-adjusted 25(OH)D (ng/ml)	1449	22.73 (7.86)	1446	22.75 (7.97)
$1,25(OH)_2D$ (pg/ml)	872	42.03 (18.05)	925	40.85 (18.42)
season-adjusted $1,25(OH)_2D$ (pg/ml)	872	42.06 (18.02)	925	40.71 (18.24)
Age (years)	1449	62.38 (5.02)	1446	62.59 (5.00)
BMI (kg/m^2)	1067	27.35 (3.94)	1177	27.17 (3.62)
PSA (ng/ml)	1449	1.39 (1.27)	1446	9.39 (26.46)
IGF-I (ng/ml)	1082	165.46 (54.49)	1405	157.69 (52.61)
IGF-II (ng/ml)	1072	759.93 (270.14)	1366	851.38 (314.96)
IGFBP-2 (ng/ml)	1069	694.08 (410.05)	1404	716.26 (407.67)
Categorical variables	**N**	**%**	**N**	**%**
[a]**Social class**				
Managerial and professional	608	41.96	594	41.08
Intermediate	235	16.22	219	15.15
Working	586	40.44	587	40.59
Missing	20	1.38	46	3.18
[b]**Family history of BPH**				
No	62	4.28	54	3.73
Yes	1272	87.78	1297	89.70
Possible	84	5.80	63	4.36
Not known/Not given	31	2.14	32	2.21
[c]**Family history of PCa**				
No	1234	85.16	1182	81.74
Yes	75	5.18	107	7.40
Missing	140	9.66	157	10.86
[d]**Smoking**				
Never	300	20.70	438	30.29
Ever	705	48.65	768	53.11
Missing	364	25.12	240	16.60
Ethnicity				
White	1428	98.55	1414	97.79
Others	15	1.04	17	1.18
Missing	6	0.41	15	1.04
[e]**Diabetes**				
Yes	936	64.60	1045	72.27
No	86	5.94	66	4.56
Missing	427	29.47	335	23.17
Study centre				
Sheffield	216	14.91	213	14.73
Newcastle	193	13.32	193	13.35
Bristol	132	9.11	130	8.99
Cardiff	148	10.21	147	10.17
Edinburgh	129	8.90	130	8.99
Birmingham	32	2.21	33	2.28
Leicester	265	18.29	264	18.26
Cambridge	163	11.25	162	11.20
Leeds	171	11.80	174	12.03

Table 1. Cont.

Categorical variables	N	%	N	%
[f]**Physical activity (per week)**				
None	481	33.20	550	38.04
1-2 times	355	24.50	367	25.38
3-4 times	159	10.97	184	12.72
5+ times	64	4.42	78	5.39
Missing	390	26.92	267	18.46

IGFBP-3, insulin-like growth factor binding protein-3; 25(OH)D, 25-hydroxyvitamin D; $1,25(OH)_2D$, 1,25-dihydroxyvitamin D; IGFBP-2, insulin-like growth factor binding protein-2; IGF-1, insulin-like growth factor 1; IGF-2, insulin-like growth factor 2; N, sample size; SD, standard deviation; IQR, inter quartile range; BPH, benign prostatic hyperplasia; PCa, PCa; BMI, body mass index; PSA, prostate specific antigen;
Normally distributed variables are presented as the means ± standard deviation, non-normally distributed variables are presented as medians (interquartile range)(*).
[a]Three-class social categorization from Rose and O'Reilly (1998).
[b]Family history of benign prostatic hyperplasia is categorized as 'yes', 'no', 'possible' and 'unknown'.
[c]Family history of PCa is coded yes if father or brother were diagnosed with PCa.
[d]Smoking status was dichotomized and defined as 'ever' (ex-smoker or current smoker) versus 'never smokers'.
[e]Diabetes status as diagnosed by a doctor and was categorized as "yes" or "no".
[f]Weekly exercise calculated from the frequency of participation in exercise of mild, moderate and strenuous intensity. Weekly exercise was categorized as 'None', '1-2 times', '3-4 times', '>5 times'.
All men included in the table have both 25-hydroxyvitamin D (25(OH)D) and IGFBP-3 measured.

Table 2. Association between circulating season-adjusted 25(OH)D (ng/ml) and 1,25$(OH)_2$D (ng/ml) with IGFBP-3 for cases and controls in the ProtecT study.

	25(OH)D						1,25$(OH)_2$D					
	N	**[a]SD unit change in IGFBP-3 (ng/ml) per SD increase in 25(OH)D (95% CI)**	***p* value**	**N**	**[b]SD unit change in IGFBP-3 (ng/ml) per SD increase in 25(OH)D (95% CI)**	***p* value**	**N**	**[a]SD unit change in IGFBP-3 (ng/ml) per SD increase in 1,25$(OH)_2$D (95% CI)**	***p* value**	**N**	**[b]SD unit change in IGFBP-3 (ng/ml) per SD increase in 1,25$(OH)_2$D (95% CI)**	***p* value**
Cases and controls	2437	0.09 (0.05, 0.13)	<0.001	1559	0.09 (0.04, 0.14)	<0.001	1428	0.09 (0.04, 0.14)	0.001	955	0.09 (0.03, 0.15)	0.003
Controls only	1071	0.07 (0.01, 0.14)	0.02	666	0.07 (-0.001, 0.15)	0.05	562	0.10 (0.03, 0.20)	0.01	354	0.12 (0.03, 0.22)	0.01
Cases only	1366	0.11 (0.05, 0.16)	<0.001	893	0.10 (0.04, 0.16)	0.002	866	0.08 (0.01, 0.15)	0.02	601	0.07 (-0.01, 0.15)	0.10

N, sample size; SD, standard deviation; IGFBP-3, insulin-like growth factor binding protein 3; 25(OH)D, 25-hydroxyvitamin D; 1,25$(OH)_2$D, 1,25 dihydroxyvitamin D; 95% CI, 95% confidence interval

[a]Adjusted for case-control status

[b]Adjusted for age, centre, case-control status, PSA levels, smoking status, social class, BMI, ethnicity, family history of PCa, history of benign hyperplasia, physical activity and diabetes status.

Table 3. Beta estimates of SD unit change in IGFBP-3 per SD unit increase in 25(OH)D and 1,25$(OH)_2$D based on one-sample Mendelian randomization analysis in the ProtecT study.

	25(OH)D					1,25$(OH)_2$D				
Instruments	**N**	**[a]Beta (95% CI)**	***p* value**	**[b]F-statistic**	**[c]*p* value (DWH)**	**N**	**[a]Beta (95% CI)**	***p* value**	**[b]F-statistic**	**[c]*p* value (DWH)**
rs3755967 (overall)	1207	-0.30 (-0.63, 0.04)	0.08	39.04	0.02	1195	-5.30 (-1.24, 0.18)	0.14	11.06	0.03
rs3755967 (controls)	487	-0.23 (-0.77, 0.31)	0.40	14.67	0.32	485	-0.35 (-1.24, 0.53)	0.44	6.35	0.23
rs3755967 (cases)	720	-0.31 (-0.72, 0.11)	0.15	24.85	0.04	710	-0.59 (-1.63, 0.44)	0.26	5.30	0.10
rs12785878 (overall)	1134	0.30 (-1.47, 2.07)	0.74	1.27	0.80	1124	0.44 (-1.32, 2.20)	0.62	1.39	0.71
rs12785878 (controls)	460	1.26 (-9.32, 11.83)	0.82	0.08	0.72	458	0.32 (-0.95, 1.60)	0.62	2.52	0.76
rs12785878 (cases)	674	0.23 (-1.33, 1.79)	0.77	1.58	0.86	666	1.64 (-11.47, 14.74)	0.81	0.07	0.68
rs10741657 (overall)	1207	0.03 (-0.67, 0.73)	0.93	7.59	0.91	1195	0.15 (-1.24, 1.55)	0.83	1.99	0.97
rs10741657 (controls)	487	1.12 (-0.42, 2.65)	0.15	3.57	0.04	485	2.14 (-2.35, 6.63)	0.35	0.97	0.05
rs10741657 (cases)	720	-0.76 (-2.01, 0.49)	0.23	4.10	0.08	710	-1.36 (-4.75, 2.04)	0.43	1.05	0.13
rs17216707 (overall)	1207	-0.36 (-1.69, 0.98)	0.60	2.48	0.50	1195	-1.69 (-11.40, 8.01)	0.73	0.18	0.45
rs17216707 (controls)	487	-0.42 (-1.47, 0.63)	0.43	4.22	0.35	485	27.35 (-1533.39, 1588.10)	0.97	0.001	0.36
rs17216707 (cases)	720	-0.40 (-7.84, 7.03)	0.92	0.08	0.88	710	-0.45 (-4.48, 3.58)	0.83	0.31	0.75
rs10745742 (overall)	1207	-2.53 (-17.64, 12.59)	0.74	0.13	0.35	1195	-6.92 (-103.79, 89.94)	0.89	0.02	0.31
rs10745742 (controls)	487	0.04 (-4.43, 4.50)	1.00	0.19	1.00	485	-0.10 (-3.80, 3.59)	0.95	0.31	0.90
rs10745742 (cases)	720	-7.13 (-105.48, 91.21)	0.89	0.02	0.28	710	-1.73 (-7.81, 4.35)	0.58	0.45	0.21
rs8018720 (overall)	1207	-1.35 (-3.91, 1.21)	0.30	1.76	0.06	1195	4.48 (-16.65, 25.62)	0.68	0.17	0.07
rs8018720 (controls)	487	0.77 (-0.64, 2.17)	0.29	3.01	0.21	485	2.46 (-6.73, 11.66)	0.60	0.29	0.22
rs8018720 (cases)	720	-0.43 (-1.14, 0.29)	0.24	9.24	0.11					

Table 3. Cont.

	25(OH)D					$1,25(OH)_2D$				
Instruments	**N**	**[a]Beta (95% CI)**	***p* value**	**[b]F-statistic**	**[c]*p* value (DWH)**	**N**	**[a]Beta (95% CI)**	***p* value**	**[b]F-statistic**	**[c]*p* value (DWH)**
[d]Allele score (overall)	1134	-0.32 (-0.64, -0.01)	0.04	46.76	0.007	1124	-0.61 (-1.33, 0.11)	0.10	11.72	0.01
[d]Allele score (controls)	460	-0.10 (-0.62, 0.41)	0.69	15.01	0.59	458	-0.15 (-0.99, 0.68)	0.72	6.20	0.49
[d]Allele score (cases)	674	-0.40 (-0.79, -0.02)	0.04	32.71	0.004	666	-0.87 (-2.01, 0.26)	0.13	5.96	0.02
[e]Synthesis score (overall)	1134	0.03 (-0.69, 0.76)	0.93	7.29	0.92	1124	0.15 (-1.06, 1.37)	0.81	2.66	0.96
[e]Synthesis score (controls)	460	0.97 (-0.74, 2.69)	0.27	2.44	0.14	458	0.97 (-0.59, 2.52)	0.22	2.76	2.77
[e]Synthesis score (cases)	674	-0.42 (-1.40, 0.56)	0.40	4.99	0.24	666	-0.98 (4.59, 2.64)	0.60	0.65	0.38
[f]Metabolism score (overall)	1207	-0.30 (-0.63, 0.03)	0.07	41.12	0.02	1195	-0.56 (-1.29, 0.16)	0.13	10.94	0.02
[f]Metabolism score (controls)	487	-0.25 (-0.74, 0.23)	0.31	18.14	0.23	485	-0.45 (-1.40, 0.50)	0.35	5.94	0.17
[f]Metabolism score (cases)	720	-0.31 (-0.74, 0.12)	0.16	23.50	0.05	710	-0.58 (-1.60, 0.43)	0.26	5.46	0.10

N, sample size; Beta, regression coefficient; 95% CI, 95% confidence interval; DWH, Durbin-Wu-Hausman; 25(OH)D, 25-hydroxyvitamin D; 1,25(OH)2D, 1,25 dihydroxyvitamin D.
[a]Beta refers to the SD unit change in IGFBP-3 levels per 1 SD unit increase in 25(OH)D or $1,25(OH)_2D$
[b]F statistic indicates how much of the variability in 25(OH)D is explained by each SNP or allele score (i.e., the strength of each SNP as an instrument for vitamin D levels. F>10 indicates a strong instrument.
[c]*p* value(DWH) is the *p* value for a test (the Durbin form of the Durbin-Wu Hausman test) for the difference between the estimates from linear regression (without additional adjustment) and instrumental variable analysis.
[d]Allele score: rs8018720, rs10745742, rs10741657, rs12785878, rs3755967 and rs17216707
[e]Synthesis score: rs12785878 and rs10741657
[f]Metabolism score: rs3755967 and rs1721

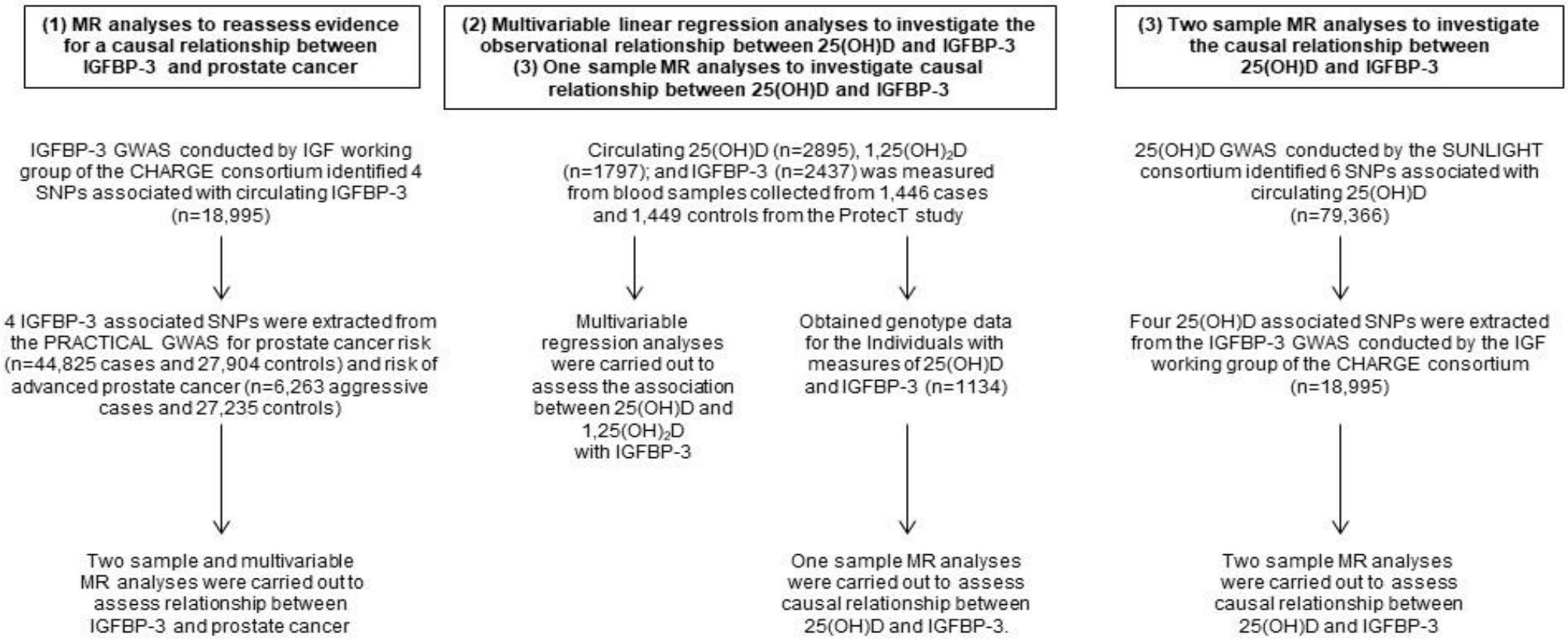

Figure 1: Flow diagram of study design

Flow diagram showing how the study population or data was selected for observational and MR analyses.

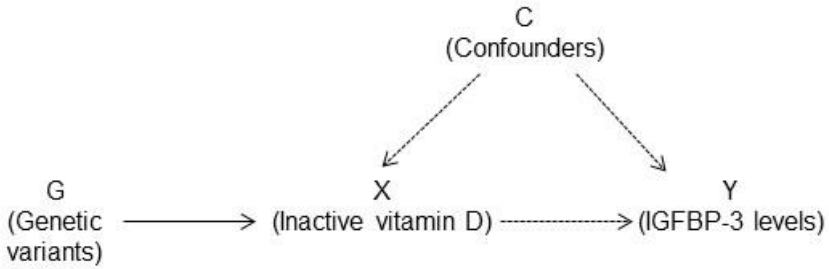

Figure 2: Mendelian randomization to infer causal nature of relationship between vitamin D and IGFBP-3
Genetic variants associated with circulating 25(OH)D levels are used as proxies for circulating vitamin D levels to assess the causal association between vitamin D and IGFBP-3. Dashed arrows show the potential sources of bias that may influence estimates derived from observational analyses or bias due to potential reverse causation. In Mendelian Randomization analyses, bias due to confounding and reverse causation is greatly reduced as the genetic variants are assigned at conception and remain unchanged throughout a person's lifetime and are not associated with confounders.

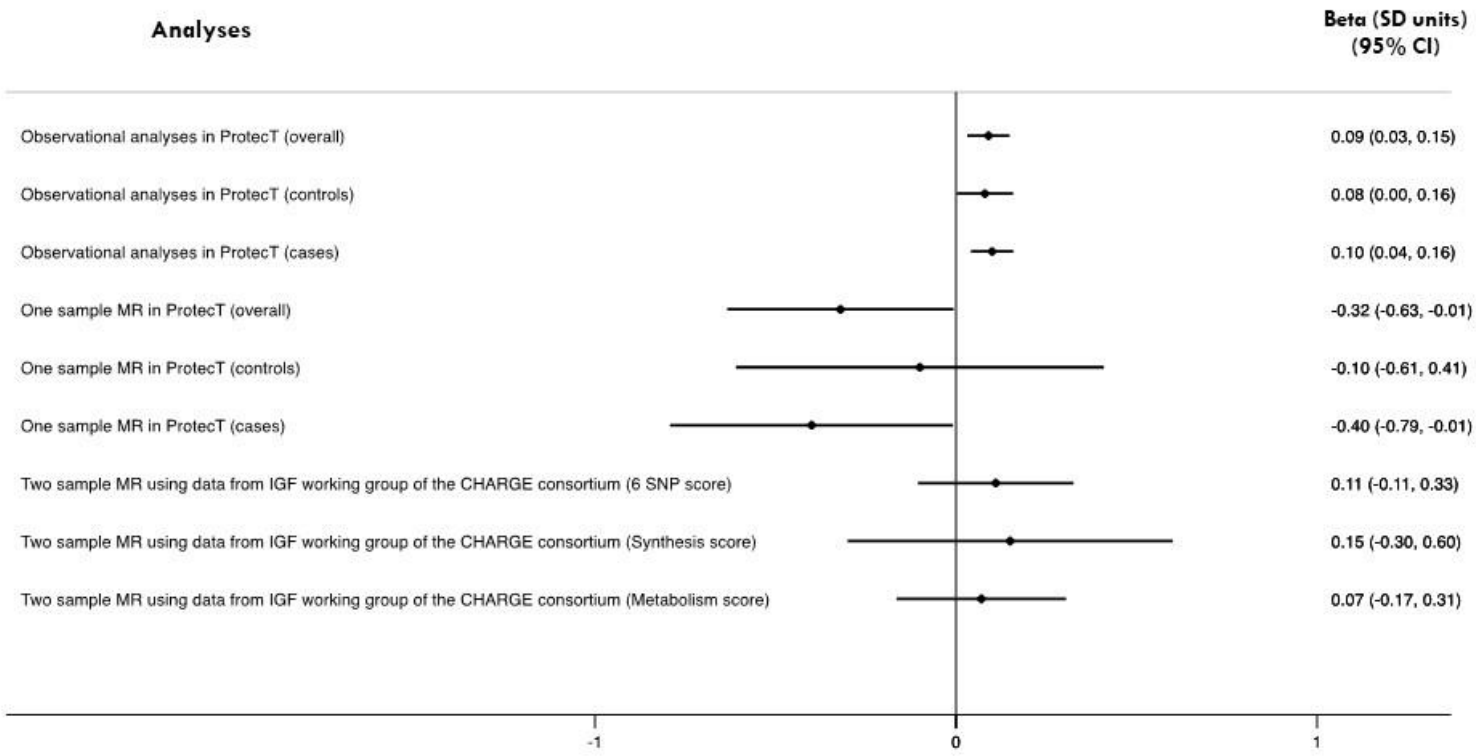

Figure 3: Forest plots of observational and instrumental variable estimates of the relationship between 25(OH)D and IGFBP-3.

The forest plot shows the estimate of the effect of 25(OH)D on IGFBP-3 from observational and MR analyses using individual level data from ProtecT or summary statistics from IGF GWAS conducted by the IGF working group of the CHARGE consortium. Diamonds represent point estimates from individual analyses. Horizontal lines represent the 95% confidence intervals.

Supplementary Table 1. Association between IGFBP-3 and prostate cancer risk and risk of advanced prostate cancer in the PRACTICAL Consortium.

PRACTICAL Consortium	Controls (N)	Cases (N)	[a]OR (95% CI)	*p* value
Risk (Two-sample MR)	27,904	44,825	IVW: 1.12 (0.91, 1.36) ML: 1.12 (1.05, 1.19)	IVW: 0.18 ML: <0.001
Risk (Multivariable MR)	27,904	44,825	1.14 (1.02, 1.28)	0.02
Advanced stage (Two-sample MR)	27,235	6,263	IVW: 1.16 (0.87, 1.55) ML: 1.17 (1.04, 1.31)	IVW: 0.19 ML:0.006
Advanced stage (Multivariable MR)	27,235	6,263	1.22 (1.07, 1.40)	0.004

MR, Mendelian Randomization; N, sample size; CI, confidence interval; IVW, inverse variance weighted; ML: maximum likelihood
[a]Associations are per 1 SD unit increase in IGFBP-3.

Supplementary Table 2. Correlation between 25(OH)D and 1,25 $(OH)_2D$ for cases and controls in ProtecT.

	1,25(OH)D		
	N	**R**	***p* value**
25(OH)D (overall)	1797	0.24	<0.001
25(OH)D (controls)	872	0.20	<0.001
25(OH)D (cases)	925	0.28	<0.001
	Season-adjusted 1,25(OH)D		
	N	**R**	***p* value**
Season-adjusted 25(OH)D (overall)	1797	0.22	<0.001
Season-adjusted 25(OH)D (controls)	872	0.18	<0.001
Season-adjusted 25(OH)D (cases)	925	0.25	<0.001

25(OH)D, 25-hydroxyvitamin D; 1,25$(OH)_2D$, 1,25-dihydroxyvitamin D; R, Pearson correlation; N, sample size

Potential Confounding factors	IGFBP-3 (ng/ml)			25(OH)D (ng/ml)			1,25(OH)$_2$D (pg/ml)		
	N	[a]Beta (95% CI)	*p* value	N	[a]Beta (95% CI)	*p* value	N	[a]Beta (95% CI)	*p* value
Age (years)	2437	-37.08 (-45.20, -28.97)	<0.001	2895	0.07 (0.01, 0.13)	0.01	1797	-0.04 (-0.21, 0.12)	0.62
Centre	2437	10.18 (-4.94, 25.30)	0.19	2895	0.06 (-0.05, 0.16)	0.28	1797	-0.005 (-0.29, 0.28)	0.97
PSA levels (ng/ml)	2437	-1.61 (-3.60, 0.37)	0.11	2895	-0.02 (-0.03, -0.0001)	0.05	1797	0.006 (-0.03, 0.04)	0.74
BMI (kg/m^2)	1904	11.19 (-1.42, 23.81)	0.08	2244	-0.20 (-0.28, -0.11)	<0.001	1438	-0.57 (-0.81, -0.32)	<0.001
[b]Smoking status	1941	-75.64 (-171.68, 20.41)	0.12	2291	-0.89 (-1.55, -0.22)	0.01	1471	-2.50 (-4.46, -0.54)	0.01
[c]Social Class	2382	-35.79 (-81.37, 9.78)	0.12	2829	-0.14 (-0.45, 0.18)	0.40	1770	-0.39 (-1.31, 0.54)	0.42
[d]Physical activity	1896	2.33 (-48.47, 53.14)	0.93	2238	1.03 (0.68, 1.38)	<0.001	1439	1.00 (-0.02, 2.03)	0.06
[e]History of BPH	2437	32.50 (-65.85, 130.86)	0.52	2895	-0.39 (-1.07, 0.30)	0.27	1797	-0.52 (-2.53, 1.49)	0.61
[f]Family history of PCa	2184	-104.02 (-270.53, 62.49)	0.22	2598	-1.09 (-2.27, 0.09)	0.07	1618	-5.10 (-8.60, -1.61)	0.004
Ethnicity	2419	-31.69 (-123.39, 60.02)	0.50	2874	-0.98 (-1.55, -0.41)	0.001	1787	-0.55 (-2.21, 1.12)	0.52
[g]Diabetes	1806	-520.11 (-706.33, -333.90)	<0.001	2133	-1.67 (-2.96, -0.39)	0.01	1371	-8.40 (-12.03, -4.78)	<0.001

Supplementary Table 3. Association between circulating 25(OH)D and IGFBP-3 levels with potential confounding factors for cases and controls in the ProtecT study.

25(OH)D, 25-hydroxyvitamin D; 1,25(OH)$_2$D, 1,25 dihydroxyvitamin D; IGFBP-3, insulin-like growth factor binding protein-3; N, sample size; Beta, regression coefficient; 95% CI, 95% confidence interval; PSA levels, prostate specific antigen levels; BMI, body mass index; BPH, benign prostatic hyperplasia; PCa, Prostate cancer

[a]Beta refers to unit change in 25(OH)D (ng/ml), 1,25(OH)$_2$D (ng/ml) or IGFBP-3 (ng/ml) levels per unit increase in confounders. Calculated using unadjusted linear regression

[b]Smoking status was dichotomized and defined as 'ever' (ex-smoker or current smoker) versus 'never smokers'.

[c]Three-class social categorization from Rose and O'Reilly (1998).

[d]Weekly exercise calculated from the frequency of participation in exercise of mild, moderate and strenuous intensity. Weekly exercise was categorized as 'None', '1-2 times', '3-4 times', '>5 times'.

[e]Family history of benign prostatic hyperplasia is categorized as 'yes', 'no', 'possible' and 'unknown'.

[f]Family history of prostate cancer is coded yes if father or brother were diagnosed with prostate cancer.[1]

[g]Diabetes status as diagnosed by a doctor and was categorized as "yes" or "no".

					25(OH)D (SD units)			1,25(OH)$_2$D (SD units)		
Genotype	**Gene**	**Chr**	**Effect/ Other allele**		**N**	**[a]Beta (95%CI)**	***p* value**	**N**	**[a]Beta (95%CI)**	***p* value**
rs3755967	*GC*	4	C/T	Overall	1501	0.29 (0.21, 0.37)	<0.001	1487	0.14 (0.06, 0.22)	0.001
				Controls	739	0.30 (0.18, 0.41)	<0.001	737	0.14 (0.03, 0.25)	0.02
				Cases	762	0.29 (0.18, 0.40)	<0.001	750	0.14 (0.03, 0.25)	0.01
rs12785878	*DHCR7*	11	T/G	Overall	1416	0.04 (-0.05, 0.14)	0.37	1404	0.06 (-0.02, 0.16)	0.20
				Controls	702	0.02 (-0.13, 0.16)	0.82	700	0.12 (-0.02, 0.25)	0.09
				Cases	714	0.07 (-0.06, 0.20)	0.29	704	0.01 (-0.12, 0.14)	0.90
rs10741657	*CYP2R1*	11	A/G	Overall	1501	0.12 (0.05, 0.19)	0.002	1487	0.03 (-0.04, 0.10)	0.42
				Controls	739	0.11 (0.003, 0.22)	0.04	737	-0.01 (-0.11, 0.10)	0.92
				Cases	762	0.13 (0.03, 0.23)	0.01	750	0.06 (-0.04, 0.16)	0.22
rs17216707	*CYP24A1*	20	T/C	Overall	1501	0.08 (-0.02, 0.17)	0.11	1487	0.02 (-0.08, 0.11)	0.74
				Controls	739	0.12 (-0.01, 0.26)	0.07	737	0.002 (-0.13, 0.13)	0.98
				Cases	762	0.03 (-0.11, 0.17)	0.66	750	0.03 (-0.10, 0.16)	0.65
rs10745742	*AMDHD1*	12	T/C	Overall	1501	0.12 (0.05, 0.19)	0.002	1487	0.03 (-0.04, 0.10)	0.42

					Controls	739	0.11 (0.003, 0.22)	0.04	737	-0.01 (-0.11, 0.10)	0.92
					Cases	762	0.13 (0.03, 0.23)	0.01	750	0.06 (-0.04, 0.16)	0.22
	rs8018720	*SEC23A*	14	G/C	Overall	1501	0.05 (-0.04, 0.15)	0.28	1487	0.02 (-0.07, 0.12)	0.60
					Controls	739	-0.11 (-0.25, 0.02)	0.11	737	0.03 (-0.10, 0.17)	0.63
					Cases	762	0.20 (0.08, 0.33)	0.002	750	0.02 (-0.11, 0.14)	0.80
	[b]Allele score				Overall	1416	0.11 (0.08, 0.15)	<0.001	1404	0.05 (0.01, 0.08)	0.01
	[b]Allele score				Controls	702	0.10 (0.02, 0.15)	<0.001	700	0.04 (-0.01, 0.09)	0.12
	[b]Allele score				Cases	714	0.13 (0.08, 0.18)	<0.001	704	0.06 (0.01, 0.10)	0.02
	[c]Synthesis score				Overall	1416	0.09 (0.03, 0.15)	0.003	1404	0.04 (-0.02, 0.09)	0.22
	[c]Synthesis score				Controls	702	0.08 (-0.01, 0.17)	0.09	700	0.04 (-0.05, 0.12)	0.41
	[c]Synthesis score				Cases	714	0.10 (0.02, 0.18)	0.01	704	0.04 (-0.03, 0.11)	0.36
	[d]Metabolism score				Overall	1416	0.20 (0.14, 0.26)	<0.001	1487	0.09 (0.03, 0.14)	0.005
	[d]Metabolism score				Controls	702	0.23 (0.14, 0.31)	<0.001	737	0.08 (-0.004, 0.17)	0.06
	[d]Metabolism score				Cases	714	0.18 0.09, 0.26)	<0.001	750	0.09 (0.01, 0.17)	0.03

Supplementary Table 4. Association between vitamin D genotypes and circulating 25(OH)D and 1,25$(OH)_2$D for cases and controls in the ProtecT study.

Chr, Chromosome; 25(OH)D, 25-hydroxyvitamin D, 1,25$(OH)_2$D, 1,25-dihydroxyvitamin D; N, sample size; Beta, regression coefficient; 95% CI, 95% confidence interval.

copy [a]Beta refers to change in 25(OH)D (SD units) levels per each additional of the effect allele. Calculated using linear regression with adjustment for case-control status for overall analyses. An SD unit of 25(OH)D is 7.8 ng/ml.

[b]Allele score: rs8018720, rs10745742, rs10741657, rs12785878, rs3755967 and rs17216707

[c]Synthesis score: rs12785878 and rs10741657

[d]Metabolism score: (rs3755967 and rs17216707)

Supplementary Table 5. Association between vitamin D genotypes and potential confounders for cases and controls in the ProtecT study.

Potential Confounders with Genotypes or Allele Score	N	[a]Beta/[b]OR (95% CI)	*p* value
[a]Age (years)			
rs3755967	1501	-0.06 (-0.45, 0.34)	0.77
rs12785878	1416	0.22 (-0.25, 0.69)	0.36
rs10741657	1501	-0.29 (-0.65, 0.07)	0.12
rs17216707	1501	0.12 (-0.34, 0.58)	0.61
rs10745742	1501	0.03 (-0.35, 0.41)	0.88
rs8018720	1501	-0.01 (-0.47, 0.45)	0.98
[c]Allele Score	1416	-0.09 (-0.26, 0.10)	0.37
[d]Synthesis score	1416	-0.10 (-0.39, 0.19)	0.51
[e]Metabolism score	1501	0.02 (-0.28, 0.31)	0.92

Potential Confounders with Genotypes or Allele Score	N	[a]Beta/[b]OR (95% CI)	*p* value
[a]Center			
rs3755967	1501	-0.06 (-0.29, 0.17)	0.61
rs12785878	1416	-0.13 (-0.40, 0.14)	0.35
rs10741657	1501	0.14 (-0.07, 0.35)	0.18
rs17216707	1501	-0.27 (-0.54, -0.003)	0.05
rs10745742	1501	0.14 (-0.08, 0.36)	0.20
rs8018720	1501	0.002 (-0.26, 0.27)	0.99
[c]Allele Score	1416	-0.001 (-0.10, 0.10)	0.99
[d]Synthesis score	1416	0.06 (-0.11, 0.23)	0.49
[e]Metabolism score	1501	-0.15 (-0.32, 0.03)	0.10
[a]PSA levels (ng/ml)			
rs3755967	1501	-0.06 (-0.15, 0.03)	0.19
rs12785878	1416	-0.08 (-0.19, 0.02)	0.13
rs10741657	1501	0.002 (-0.08, 0.08)	0.97
rs17216707	1501	0.09 (-0.02, 0.19)	0.10
rs10745742	1501	-0.08 (-0.16, 0.01)	0.07
rs8018720	1501	0.03 (-0.07, 0.13)	0.56
[c]Allele Score	1416	-0.02 (-0.06, 0.02)	0.28
[d]Synthesis score	1416	-0.02 (-0.09, 0.04)	0.48
[e]Metabolism score	1501	0.002 (-0.06, 0.07)	0.94

Supplementary Table 5 Cont.

Potential Confounders with Genotypes or Allele Score	N	Beta[a]/OR[b] (95% CI)	*p* value
[a]BMI (kg/m^2)			
rs3755967	1196	0.13 (-0.20, 0.45)	0.45
rs12785878	1130	0.15 (-0.24, 0.54)	0.45
rs10741657	1196	-0.17 (-0.46, 0.13)	0.27
rs17216707	1196	0.003 (-0.38, 0.38)	0.99
rs10745742	1196	0.02 (-0.29, 0.33)	0.89
rs8018720	1196	-0.03 (-0.41, 0.34)	0.86
[c]Allele Score	1130	0.02 (-0.13, 0.16)	0.81
[d]Synthesis score	1130	-0.05 (-0.29, 0.19)	0.69
[e]Metabolism score	1196	0.07 (-0.17, 0.31)	0.56
[b]Smoking status			
rs3755967	1226	1.03 (0.86, 1.23)	0.77
rs12785878	1157	1.08 (0.87, 1.34)	0.47
rs10741657	1226	1.12 (0.95, 1.32)	0.17
rs17216707	1226	0.99 (0.80, 1.23)	0.95
rs10745742	1226	1.12 (0.94, 1.33)	0.20
rs8018720	1226	1.02 (0.83, 1.26)	0.85
[c]Allele Score	1157	1.06 (0.98, 1.15)	0.16
[d]Synthesis score	1157	1.07 (0.94, 1.22)	0.29
[e]Metabolism score	1226	1.01 (0.89, 1.16)	0.86
[a]Social class			
rs3755967	1476	-0.02 (-0.09, 0.05)	0.57
rs12785878	1396	0.008 (-0.08, 0.09)	0.85
rs10741657	1476	0.02 (-0.05, 0.08)	0.62
rs17216707	1476	0.08 (-0.007, 0.16)	0.07
rs10745742	1476	-0.03 (-0.10, 0.03)	0.33
rs8018720	1476	-0.03 (-0.12, 0.05)	0.46
[c]Allele Score	1396	0.003 (-0.03, 0.04)	0.84

[d]Synthesis score	1396	0.02 (-0.03, 0.07)	0.51
[e]Metabolism score	1476	0.02 (-0.03, 0.07)	0.47
[a]Physical activity			
rs3755967	1196	-0.04 (-0.12, 0.04)	0.29
rs12785878	1128	-0.03 (-0.13, 0.06)	0.53
rs10741657	1196	-0.002 (-0.07, 0.07)	0.97
rs17216707	1196	-0.06 (-0.16, 0.03)	0.19
rs10745742	1196	-0.03 (-0.10, 0.05)	0.48
rs8018720	1196	0.008 (-0.09, 0.10)	0.86
[c]Allele Score	1128	-0.02 (-0.06, 0.01)	0.20
[d]Synthesis score	1128	-0.009 (-0.07, 0.05)	0.76
[e]Metabolism score	1196	-0.05 (-0.11, 0.01)	0.10

Supplementary Table 5 Cont.

Potential Confounders with Genotypes or Allele Score	N	Beta[a]/OR[b] (95% CI)	*p* value
[a]History of BPH			
rs3755967	1501	-0.009 (-0.04, 0.03)	0.62
rs12785878	1416	0.002 (-0.04, 0.04)	0.94
rs10741657	1501	0.01 (-0.02, 0.04)	0.41
rs17216707	1501	-0.02 (-0.06, 0.02)	0.39
rs10745742	1501	0.01 (-0.02, 0.04)	0.49
rs8018720	1501	-0.002 (-0.04, 0.04)	0.89
[c]Allele Score	1416	0.0004 (-0.01, 0.02)	0.96
[d]Synthesis score	1416	0.01 (-0.02, 0.03)	0.44
[e]Metabolism score	1501	-0.01 (-0.04, 0.01)	0.35
[b]Family history of PCa			
rs3755967	1350	0.94 (0.69, 1.29)	0.72
rs12785878	1277	0.90 (0.62, 1.30)	0.58
rs10741657	1350	1.15 (0.86, 1.53)	0.35
rs17216707	1350	0.81 (0.57, 1.15)	0.23
rs10745742	1350	0.70 (0.51, 0.96)	0.03
rs8018720	1350	0.98 (0.68, 1.42)	0.92

[c]Allele Score	1277	0.90 (0.78, 1.03)	0.13
[d]Synthesis score	1277	1.01 (0.80, 1.27)	0.93
[e]Metabolism score	1350	0.88 (0.70, 1.11)	0.30
[b]Diabetes			
rs3755967	1141	1.00 (0.71, 1.40)	1.00
rs12785878	1074	1.05 (0.70, 1.57)	0.81
rs10741657	1141	1.05 (0.78, 1.43)	0.73
rs17216707	1141	0.90 (0.61, 1.32)	0.58
rs10745742	1141	1.25 (0.91, 1.72)	0.16
rs8018720	1141	1.47 (1.02, 2.11)	0.04
[c]Allele Score	1074	1.14 (0.98, 1.32)	0.09
[d]Synthesis score	1074	1.10 (0.86, 1.40)	0.45
[e]Metabolism score	1141	0.96 (0.74, 1.23)	0.72

N, sample size; Beta, regression coefficient; OR, odds ratio; 95% CI, 95% confidence interval; PSA levels, prostate specific antigen levels; BMI, body mass index; BPH, benign prostatic hyperplasia; PCa, Prostate

cancer.
[a]Per-allele effects were obtained by linear regression for continuous variables.
[b]Per-allele effects were obtained by logistic regression for binary variables.
[d] Allele score: rs8018720, rs10745742, rs10741657, rs12785878, rs3755967 and rs17216707.
[e]Synthesis score: rs12785878 and rs10741657.
[f] Metabolism score: rs3755967 and rs17216707.

Supplementary Table 6. Leave one out analyses repeated for each possible combination of 25-hydroxyvitamin D genetic instruments

Instruments	Beta (95% CI)	*p*
Six 25(OH)D associated SNPs excluding rs10741657	0.11 (-0.11, 0.33)	0.34
Six 25(OH)D associated SNPs excluding rs10745742	0.11 (-0.11, 0.32)	0.32
Six 25(OH)D associated SNPs excluding rs12785878	0.09 (-0.13, 0.31)	0.41
Six 25(OH)D associated SNPs excluding rs17216707	0.11 (-0.10, 0.33)	0.29
Six 25(OH)D associated SNPs excluding rs3755967	0.17 (-0.20, 0.55)	0.37
Six 25(OH)D associated SNPs excluding rs8018720	0.09 (-0.12, 0.30)	0.42
Six 25(OH)D associated SNPs excluding rs10741657	0.11 (-0.10, 0.31)	0.32

25(OH)D, 25-hydroxyvitamin D; SNP, single nucleotide polymorphisms; CI, confidence interval; *p*, p-value.

Supplementary Table 7. Causal effect of IGFBP-3 on 25(OH)D using two sample MR analysis.

Instruments	Beta (95% CI)	p-value
rs1065656	0.01 (-0.03, 0.05)	0.57
rs11977526	0.004 (-0.01, 0.02)	0.62
rs4234798	0.05 (0.01, 0.09)	0.03
rs700753	0.03 (0.01, 0.004)	0.02
Overall (IVW)	0.01 (-0.003, 0.03)	0.10

25(OH)D, 25-hydroxyvitamin D; SNP, single nucleotide polymorphisms; CI, confidence interval; IVW, inverse variance weighted

PIs from the PRACTICAL consortium:

(Information of the consortium can be found at http://PRACTICAL.icr.ac.uk)

Rosalind A. Eeles[1,2], Christopher A. Haiman[3], ZSofia Kote-Jarai[1], Fredrick R. Schumacher[4,5], Sara Benlloch[6,1], Ali Amin Al Olama[6,9], Muir Kenneth[,8], Sonja I. Berndt[10], David V. Conti[3], Fredrik Wiklund[11], Stephen Chanock[10], Victoria L. Stevens[12], Catherine M. Tangen[13], Jyotsna Batra[14,15], Judith Clements[14,15], APCB BioResource[14], Henrik Gronberg[11], Nora Pashayan[16,17], Johanna Schleutker[18,19,20], Demetrius Albanes[10], Stephanie Weinstein[10], Alicja Wolk[22], Catharine West[23], Lorelei Mucci[24], Géraldine Cancel-Tassin[25,26], Stella Koutros[10], Karina Dalsgaard Sorensen[27,28], Lovise Maehle[29], David E. Neal[30,31], Freddie C. Hamdy[32], Jenny L. Donovan[33], Ruth C. Travis[34], Robert J. Hamilton[35], Sue Ann Ingles[3], Barry Rosenstein[36,37], Yong-Jie Lu[38], Graham G. Giles[39,40], Adam S. Kibel[41], Ana Vega[42], Manolis Kogevinas[43,44,45,46], Kathryn L. Penney[47], Jong Y. Park[48], Janet L. Stanford[49,50], Cezary Cybulski[51], Børge G. Nordestgaard[52,53], Hermann Brenner[54,55,56], Christiane Maier[57], Jeri Kim[58], Esther M. John[59,60], Manuel R. Teixeira[61,62], Susan L. Neuhausen[63], Kim De Ruyck[64], Azad Razack[65], Lisa F. Newcomb[49,66], Davor Lessel[67], Radka Kaneva[68], Nawaid Usmani[69,70], Frank Claessens[71], Paul A. Townsend[72], Manuela Gago Dominguez[73,74], Monique J. Roobol[75], Florence Menegaux[76], Kay-Tee Khaw[77], Lisa Cannon-Albright[78,79], Hardev Pandha[80], Stephen N. Thibodeau[81]

[1] The Institute of Cancer Research, London, UK.

[2] Royal Marsden NHS Foundation Trust, London, UK.

[3] Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA, USA.

[4] Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA.

[5] Seidman Cancer Center, University Hospitals, Cleveland, OH, USA.

[6] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK.

[7] Institute of Population Health, University of Manchester, Manchester, UK.

[8] Warwick Medical School, University of Warwick, Coventry, UK.

[9] University of Cambridge, Department of Clinical Neurosciences, Cambridge, UK.

[10] Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD, USA.

[11] Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden.

[12] Epidemiology Research Program, American Cancer Society, 250 Williams Street, Atlanta, GA, USA.

[13] SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

[14] Australian Prostate Cancer Research Centre-Qld, Institute of Health and Biomedical Innovation and School of Biomedical Science, Queensland University of Technology, Brisbane, Queensland, Australia.

[15] Translational Research Institute, Brisbane, Queensland, Australia.

[16] University College London, Department of Applied Health Research, London, UK.

[17] Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Strangeways Laboratory, Cambridge, UK.

[18] Department of Medical Biochemistry and Genetics, Institute of Biomedicine, University of Turku, Finland.

[19] Tyks Microbiology and Genetics, Department of Medical Genetics, Turku University Hospital, Finland.

[20] BioMediTech, University of Tampere, Tampere, Finland.

[21] Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD, USA.

[22] Division of Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Sweden.

[23] Institute of Cancer Sciences, University of Manchester, Manchester Academic Health Science Centre, Radiotherapy Related Research, The Christie Hospital NHS Foundation Trust, Manchester, UK.

[24] Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA.

[25] CeRePP, Pitie-Salpetriere Hospital, Paris, France.

[26] UPMC Univ Paris 06, GRC N°5 ONCOTYPE-URO, CeRePP, Tenon Hospital, Paris, France.

[27] Department of Molecular Medicine, Aarhus University Hospital, Denmark.

[28] Department of Clinical Medicine, Aarhus University, Denmark.

[29] Department of Medical Genetics, Oslo University Hospital, Norway.

[30] University of Cambridge, Department of Oncology, Addenbrooke's Hospital, Cambridge, UK.

[31] Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge, UK.

[32] Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK, Faculty of Medical Science, University of Oxford, John Radcliffe Hospital, Oxford, UK.

[33] School of Social and Community Medicine, University of Bristol, Bristol, UK.

[34] Cancer Epidemiology, Nuffield Department of Population Health University of Oxford, Oxford, UK.

[35] Dept. of Surgical Oncology, Princess Margaret Cancer Centre, Toronto, Canada.

[36] Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[37] Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[38] Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary University of London, John Vane Science Centre, London, UK.

[39] Cancer Epidemiology and Intelligence Division, The Cancer Council Victoria, Melbourne, Victoria, Australia.

[40] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia.

[41] Division of Urologic Surgery, Brigham and Womens Hospital, Boston, MA, USA.

[42] Fundación Pública Galega de Medicina Xenómica-SERGAS, Grupo de Medicina Xenómica, CIBERER, IDIS, Santiago de Compostela, Spain.

[43] Centre for Research in Environmental Epidemiology (CREAL), Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain.

[44] CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain.

[45] IMIM (Hospital del Mar Research Institute), Barcelona, Spain.

[46] Universitat Pompeu Fabra (UPF), Barcelona, Spain.

[47] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital/Harvard Medical School, Boston, MA, USA.

[48] Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, USA.

[49] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

[50] Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington, USA.

[51] International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland.

[52] Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

[53] Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark.

[54] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany.

[55] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany.

[56] Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany.

[57] Institute for Human Genetics, University Hospital Ulm, Ulm, Germany.

[58] The University of Texas M. D. Anderson Cancer Center, Department of Genitourinary Medical Oncology, Houston, TX, USA.

[59] Cancer Prevention Institute of California, Fremont, CA, USA.

[60] Department of Health Research & Policy (Epidemiology) and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA , USA.

[61] Department of Genetics, Portuguese Oncology Institute of Porto, Porto, Portugal.

[62] Biomedical Sciences Institute (ICBAS), University of Porto, Porto, Portugal.

[63] Department of Population Sciences, Beckman Research Institute of the City of Hope, Duarte, CA, USA.

[64] Ghent University, Faculty of Medicine and Health Sciences, Basic Medical Sciences, Gent, Belgium.

[65] Department of Surgery, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia.

[66] Department of Urology, University of Washington, Seattle, WA, USA.

[67] Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

[68] Molecular Medicine Center, Department of Medical Chemistry and Biochemistry, Medical University, Sofia, Bulgaria.

[69] Department of Oncology, Cross Cancer Institute, University of Alberta, Edmonton, Alberta, Canada.

[70] Division of Radiation Oncology, Cross Cancer Institute, Edmonton, Alberta, Canada.

[71] Molecular Endocrinology Laboratory, Department of Cellular and Molecular Medicine, KU Leuven, Leuven, Belgium.

[72] Institute of Cancer Sciences, Manchester Cancer Research Centre, University of Manchester, Manchester Academic Health Science Centre, St Mary's Hospital, Manchester, UK.

[73] Genomic Medicine Group, Galician Foundation of Genomic Medicine, Instituto de Investigacion Sanitaria de Santiago de Compostela (IDIS), Complejo Hospitalario Universitario de Santiago, Servicio Galego de Saúde, SERGAS, Santiago De Compostela, Spain.

[74] University of California San Diego, Moores Cancer Center, La Jolla, CA, USA.

[75] Department of Urology, Erasmus University Medical Center, Rotterdam, the Netherlands.

[76] Cancer & Environment Group, Center for Research in Epidemiology and Population Health (CESP), INSERM, University Paris-Sud, University Paris-Saclay, Villejuif, France.

[77] Clinical Gerontology Unit, University of Cambridge, Cambridge, UK.

[78] Division of Genetic Epidemiology, Department of Medicine, University of Utah School of Medicine, Salt Lake City, Utah, USA.

[79] George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT, USA.

[80] The University of Surrey, Guildford, Surrey, UK.

[81] Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA.

Funding Acknowledgements

The PRACTICAL consortium was supported by Cancer Research UK Grants C5047/A7357, C1287/A10118, C1287/A16563, C5047/A3354, C5047/A10692, C16913/A6135, European Commission's Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175), and The National Institute of Health (NIH) Cancer Post-Cancer GWAS initiative grant: No. 1 U19 CA 148537-01 (the GAME-ON initiative). Genotyping of the OncoArray was funded by the US National Institutes of Health (NIH) [U19 CA 148537 for ELucidating Loci Involved in Prostate cancer SuscEptibility (ELLIPSE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN268201200008I]. Additional analytic support was provided by NIH NCI U01 CA188392 (PI: Schumacher). Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund. We would also like to thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK (now Prostate Action), The Orchid Cancer Appeal, The National Cancer Research Network UK, The National Cancer Research Institute (NCRI) UK. We are grateful for support of NIHR funding to the NIHR Biomedical Research Centre at The Institute of Cancer Research and The Royal Marsden NHS Foundation Trust

Supplementary Methods and Tables

Study populations and data sources for observational and MR analyses

ProtecT study

Individual level analyses were based within the ProtecT trial (ISRCTN20141297), a large population-based cohort of approximately 80,000 men aged 50-69 years from across the UK who underwent PSA testing for prostate cancer, within which are nested case-control studies (1, 2). All men without evidence of prostate cancer (PSA test <3ng/mL, or men with PSA >3ng/mL but had at least one negative diagnostic biopsy) were eligible for selection as controls. Controls were randomly selected from the same stratum (5-year age band and GP/family practice) as the cases who had provided a non-fasted blood sample at the prostate check clinic. Circulating concentrations of IGFBP-3 were measured using an in-house radioimmunoassay from blood serum. The intra-class correlations (ICC) for within-assay variability and between-assay variability for IGFBP-3

were 0.88 and 0.71, respectively. Circulating $25(OH)D_2$ and $25(OH)D_3$ in blood plasma (collected at diagnosis) were measured using tandem mass spectrometry. Circulating concentrations of $25(OH)D_2$ and $25(OH)D_3$ were measured in nanograms per millilitre (ng/mL) where 1 ng/mL = 2.5 nmol/L (nanomoles per litre). Total 25(OH)D (ng/mL) was calculated as the summation of $25(OH)D_2$ and $25(OH)D_3$. Circulating concentrations of $1,25(OH)_2D$ (pg/mL) were measured in blood plasma using immunoassay (single batch of reagents) over a 2-month period, as previously described (3). Circulating concentrations of $1,25(OH)_2D$ was measured in picomoles per liter (pmol/L) where 1 pg/mL = 2.6 pmol/L. Measures of height, weight, weekly exercise, smoking status, alcohol consumption, family history of prostate cancer (father and brother), history of benign prostatic hyperplasia (BPH), diabetes, occupational social class and self-reported ethnicity were collected before receipt of the initial PSA test result, either by questionnaire or by nurse interview, as previously described (4). Genome-wide genotyping of participants was carried out using the Illumina Human660W-Quad_v1_A array (Illumina Inc., SanDiego, CA). After genotyping, quality control (QC) process was performed before imputation. We excluded individuals on the basis of the following: sex mismatches, minimal (<0.325) or excessive heterozygosity (>0.345), disproportionate levels of individual missingness (>3 %), cryptic relatedness measured as proportion of identity by descent (IBD > 0.1) and insufficient sample replication (IBD < 0.8). All individuals with non-European ancestry and SNPs with a minor allele frequency (MAF) below 1%, a call rate of <95% or out of Hardy-Weinberg equilibrium ($p<5 \times 10^{-7}$) were removed (5). All men provided written informed consent prior to inclusion in the study. Trent Multicentre Research Ethics Committee (MREC) approved the ProtecT study (MREC/01/4/025) and the associated ProMPT study which collected biological material (MREC/01/4/061).

PRACTICAL Consortium (PRostate cancer AssoCiation group To Investigate Cancer Associated aLterations in the genome)

The PRACTICAL consortium conducted a GWAS meta-analysis to identify SNPs associated with prostate cancer risk in analyses involving 44,825 prostate cancer cases and 27,904 controls of European ancestry. They also conducted a GWAS to identify SNPs associated with risk of advanced prostate cancer in analyses involving 6,263 aggressive prostate cancer cases (defined as Gleason ≥8, prostate-specific antigen (PSA) >100ng/mL, metastatic disease (M1) or death from prostate cancer) and 27,235 controls (summary data on high-grade prostate cancer alone was not available). Full details of the individual participating studies have been published previously (6). Genotyping of PRACTICAL samples were performed using an Illumina Custom Infinium genotyping array (OncoArray) consisting of ~570,000 single nucleotide polymorphisms (SNPs) and imputed using the 1000 Genomes Project as a reference panel as previously described (7). All SNPs with poor imputation quality ($r^2<0.30$), a minor allele frequency of <1%, a call rate of <98% or evidence of Hardy-Weinberg equilibrium ($p<10^{-7}$ in controls or $p<10^{-12}$ in cases) were removed. All studies in PRACTICAL have the relevant Institutional Review Board approval from each country, in accordance with the Declaration of Helsinki.

IGF working group of the CHARGE consortium

The IGF working group of the CHARGE consortium included 13 studies that participated in the GWAS for IGFBP-3 involving 18,995 participants (8,053 men and 10,942 women) of European ancestry (8). This GWAS meta-analysis identified four SNPs (rs11977526, rs700753, rs1065656, rs4234798) that had genome-wide significant associations with IGFBP-3. The methods applied to measure IGFBP-3 levels differed among the participating cohorts and include enzyme-linked immunosorbent assays (ELISA), sandwich ELISA, coated tube immunoradiometric assays, chemiluminescent immunometric assays, radioimmunoassays (RIA) and automated two-site chemiluminescence immunoassay, as previously described (8). Participants were genotyped on genome-wide arrays and SNPs were imputed using the HapMap2 reference panel. Association analyses in individual studies were performed on IGFBP-3 levels measured in ng/mL using a multiple linear regression with an additive genetic model based on allele dosages adjusted for age and stratified by sex. All cohorts accounted for relatedness, population substructure using genetic principal components, study centre and laboratory batch of IGFBP-3 measurement where applicable. Individuals of non-European ancestry, with missing phenotype data, a diagnosis of growth hormone deficiency or known use of human growth hormones were excluded from the analyses (8).

25(OH)D Genome Wide Association Study (GWAS) by the SUNLIGHT consortium

The SUNLIGHT consortium included 31 studies that participated in the GWAS for 25(OH)D involving 79,366 participants of European ancestry (9). This GWAS meta-analysis identified six SNPs (rs8018720, rs10745742, rs10741657, rs12785878, rs3755967 and rs17216707) that had genome-wide significant associations with 25(OH)D. The methods applied to measure 25(OH)D levels differed among the participating cohorts and include ELISA, high performance liquid chromatography tandem mass spectrum (HPLC), radioimmunoassay (RIA) amd electrochemiluminescence immunoassays. Participants were genotyped on genome-wide arrays and SNPs were imputed using the Haplotype Reference Consortium (HRC) panel. Association analyses in individual studies were performed on natural-log transformed 25(OH)D, using a linear regression with an additive genetic model adjusted for month of sample collection, age, sex, body mass index and principal components (capturing genetic ancestry). To combine results across the contributing cohorts, a fixed effects inverse variance weighted meta-analysis was performed with control for population structure within each cohort and quality control thresholds of MAF>0.05, imputation info score>0.8, Hardy-Weinberg equilibrium (HWE) >1x10 and a minimum of 2 studies and 10,000 individuals contributing to each reported SNP-phenotype association.

Statistical analyses

Observational relationship between circulating $1,25(OH)_2D$ and IGFBP-3 levels in ProtecT

Using data from the ProtecT study, multivariable regression analyses (with adjustment for case-control status) was employed to test observational associations between circulating $1,25(OH)_2D$ with IGFBP-3 levels. Association of $1,25(OH)_2D$ and IGFBP-3 with potential confounders and other variables (age, center, PSA levels (logged), BMI, smoking status, social class, history of benign hyperplasia, family history of prostate cancer and diabetes status) was also estimated using linear regression. Observational estimates of the SD unit change in IGFBP-3 (ng/ml) per SD unit increase in season-adjusted $1,25(OH)_2D$ (pg/ml) were estimated using linear regression with adjustment for confounders (age, center, case-control status, PSA levels (logged), smoking status, social class, family history of prostate cancer, body mass index (BMI) and diabetes status).

Causal associations between IGFBP-3 and 25(OH)D: two sample MR

In the reverse direction, we investigated the causal effect of IGFBP-3 on 25(OH)D using the two-sample MR analyses. The SNP-exposure (IGFBP-3) and SNP-outcome (25(OH)D) estimates for the four IGFBP-3 associated SNPs were combined using the IVW method(10). The IVW method (similar method to that used in the two-sample MR analysis of IGFBP-3 and PCa) is equivalent to the two-stage least square analyses using data from ProtecT(11). As only a small number of SNPs were included in our MR analyses, methods to test for pleiotropy such as MR-Egger regression(11), weighted median(12) and mode method(13) were not conducted as these lack power with small number of SNPs.

References

1. Lane A, Metcalfe C, Young GJ, Peters TJ, Blazeby J, Avery KN, et al. Patient-reported outcomes in the ProtecT randomized trial of clinically localized prostate cancer treatments: study design, and baseline urinary, bowel and sexual function and quality of life. BJU international. 2016;118(6):869-79.
2. Lane JA, Hamdy FC, Martin RM, Turner EL, Neal DE, Donovan JL. Latest results from the UK trials evaluating prostate cancer screening and treatment: the CAP and ProtecT studies. European journal of cancer. 2010;46(17):3095-101.

3. Gilbert R, Metcalfe C, Fraser WD, Donovan J, Hamdy F, Neal DE, et al. Associations of circulating retinol, vitamin E, and 1,25-dihydroxyvitamin D with prostate cancer diagnosis, stage, and grade. Cancer causes & control : CCC. 2012;23(11):1865-73.
4. Gilbert R, Metcalfe C, Fraser WD, Donovan J, Hamdy F, Neal DE, et al. Associations of circulating 25-hydroxyvitamin D with prostate cancer diagnosis, stage and grade. International journal of cancer. 2012;131(5):1187-96.
5. Bonilla C, Lewis SJ, Rowlands MA, Gaunt TR, Davey Smith G, Gunnell D, et al. Assessing the role of insulin-like growth factors and binding proteins in prostate cancer using Mendelian randomization: Genetic variants as instruments for circulating levels. International journal of cancer. 2016;139(7):1520-33.
6. Al Olama AA, Kote-Jarai Z, Berndt SI, Conti DV, Schumacher F, Han Y, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nature genetics. 2014;46(10):1103-9.
7. Amos CI, Dennis J, Wang Z, Byun J, Schumacher FR, Gayther SA, et al. The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2017;26(1):126-35.
8. Teumer A, Qi Q, Nethander M, Aschard H, Bandinelli S, Beekman M, et al. Genomewide meta-analysis identifies loci associated with IGF-I and IGFBP-3 levels with impact on age-related traits. Aging cell. 2016;15(5):811-24.
9. Jiang X, O'Reilly PF, Aschard H, Hsu YH, Richards JB, Dupuis J, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat Commun. 2018;9(1):260.
10. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic epidemiology. 2013;37(7):658-65.
11. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International journal of epidemiology. 2015;44(2):512-25.
12. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genetic epidemiology. 2016;40(4):304-14.
13. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. International journal of epidemiology. 2017;46(6):1985-98.